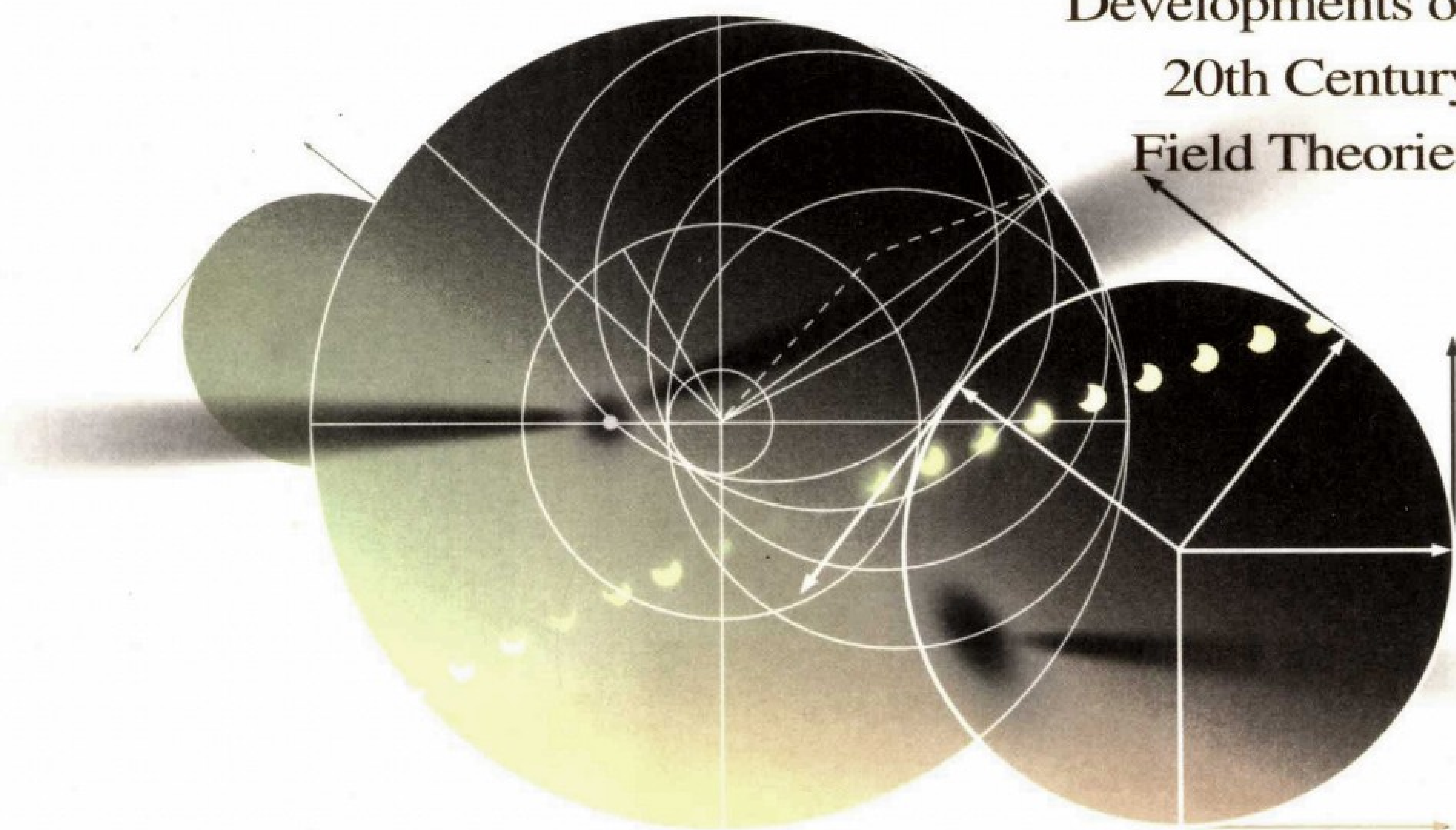


开放人文



Conceptual Developments of 20th Century Field Theories

[美] 曹天予 著　吴新忠　李宏芳　李继堂 译　桂起权 校

Tian Yu Cao

20世纪场论的概念发展

上海世纪出版集团

20世纪场论的概念发展

曹天予 著

吴新忠、李宏芳、李继堂 译 桂起权 校

何祚庥、范岱年 序

世纪出版集团 上海科技教育出版社

图书在版编目(CIP)数据

20 世纪场论的概念发展/(美)曹天予著;吴新忠,李宏芳,李继堂译.—上海:上海科技教育出版社,2008.12

(世纪人文系列丛书.开放人文)

ISBN 978-7-5428-4536-8

Ⅰ.2... Ⅱ.①曹...②吴...③李...④李... Ⅲ.场论—研究 Ⅳ.O412.3

中国版本图书馆 CIP 数据核字(2008)第 082239 号

责任编辑 郑华秀

装帧设计 陆智昌 朱赢椿

---

20 世纪场论的概念发展

[美]曹天予 著

吴新忠 李宏芳 李继堂 译 桂起权 校

出 版 世纪出版集团 上海科技教育出版社
(200235 上海冠生园路 393 号 www.ewen.cc)

发 行 上海世纪出版集团发行中心

印 刷 上海江杨印刷厂

开 本 635×965 mm 1/16

印 张 39

插 页 4

字 数 523 000

版 次 2008 年 12 月第 1 版

印 次 2008 年 12 月第 1 次印刷

ISBN 978-7-5428-4536-8/N·741

图 字 09-2004-741 号

定 价 55.00 元

Conceptual Developments of 20th Century Field Theories
by Tian Yu Cao

First published 1997
First paperback edition(with corrections) 1998



出版说明

自中西文明发生碰撞以来，百余年的中国现代文化建设即无可避免地担负起双重使命。梳理和探究西方文明的根源及脉络，已成为我们理解并提升自身要义的借镜，整理和传承中国文明的传统，更是我们实现并弘扬自身价值的根本。此二者的交汇，乃是塑造现代中国之精神品格的必由进路。世纪出版集团倾力编辑世纪人文系列丛书之宗旨亦在于此。

世纪人文系列丛书包涵“世纪文库”、“世纪前沿”、“袖珍经典”、“大学经典”及“开放人文”五个界面，各成系列，相得益彰。

“厘清西方思想脉络，更新中国学术传统”，为“世纪文库”之编辑指针。文库分为中西两大书系。中学书系由清末民初开始，全面整理中国近现代以来的学术著作，以期为今人反思现代中国的社会和精神处境铺建思考的进阶；西学书系旨在从西方文明的整体进程出发，系统译介自古希腊罗马以降的经典文献，借此展现西方思想传统的生发流变过程，从而为我们返回现代中国之核心问题奠定坚实的文本基础。与之呼应，“世纪前沿”着重关注二战以来全球范围内学术思想的重要论题与最新进展，展示各学科领域的新近成果和当代文化思潮演化的各种向度。“袖珍经典”则以相对简约的形式，收录名家大师们在体裁和风格上独具特色的经典作品，阐幽发微，意趣兼得。

遵循现代人文教育和公民教育的理念，秉承“通达民情，化育人心”的中国传统教育精神，“大学经典”依据中西文明传统的知识谱系及其价值内涵，将人类历史上具有人文内涵的经典作品编辑成为大学教育的基础读本，应时代所需，顺时势所趋，为塑造现代中国人的人文素养、公民意识和国家精神倾力尽心。“开放人文”旨在提供全景式的人文阅读平台，从文学、历史、艺术、科学等多个面向调动读者的阅读愉悦，寓学于乐，寓乐于心，为广大读者陶冶心性，培植情操。

“大学之道，在明明德，在新民，在止于至善”（《大学》）。温古知今，止于至善，是人类得以理解生命价值的人文情怀，亦是文明得以传承和发展的精神契机。欲实现中华民族的伟大复兴，必先培育中华民族的文化精神；由此，我们深知现代中国出版人的职责所在，以我之不懈努力，做一代又一代中国人的文化脊梁。

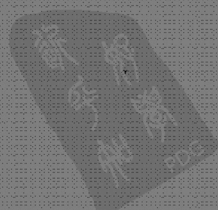
上海世纪出版集团

世纪人文系列丛书编辑委员会

2005年1月

20世纪场论的概念发展

献给罗莎（Rosa）



内 容 提 要

本书给出了20世纪场论的概念发展的广泛综述，涉及从广义相对论到量子场论和规范场论的概念发展。作者针对这些理论的概念基础给出了一种历史批判性的阐述，揭示了这些自然概念演进的模式。

理论物理学的专家与学生将在这本书中发现对他们所从事的学科的基本问题的一种论述，这有助于他们理解本学科的内在逻辑和发展动力。另外，对于职业科学史学家和科学哲学家，尤其是物理学史学家和物理学哲学家，本书将为他们对这些理论进行更深刻的历史、文化和社会学分析提供概念基础。这本书也包含相当多的哲学(形而上学的、方法论的和语义学的)反思材料。最后，有一定科学素养的一般读者将发现，本书对物理世界的当代概念作出了较之本学科的通俗说明更为深入的分析。

当代数理物理学核心部分的这种令人着迷的论述，将使从专业研究者到受过教育的普通读者的各个水平上的物理学家、科学史学家、科学哲学家和科学社会学家产生浓厚的兴趣。

作 者 简 介

曹天予(1941—)，剑桥大学科学史和科学哲学博士，美国波士顿大学哲学系教授。曾是英国剑桥大学三一学院(Trinity College, 1985—1990)和牛津大学万灵学院(All Souls College, 2000—2001)的研究员；并在美国哈佛大学(1988—1992)、麻省理工学院(1992—1994)和伦敦经济学院(1999—2000)做过访问研究；史密森学会(Smithsonian Institution, 2000 年 3—8 月)的高级研究员；普林斯顿高等研究院自然科学院和史学院(2004—2005)的成员。除英美两国外，还多次应邀到法国、德国、加拿大、俄国、瑞士等国和香港、台湾等地区讲学。主要研究方向为当代理论物理学中的概念问题及其历史演变。除本书外，还发表了有关重正化、对称性破缺和雷杰化(Reggeization，调和量子场论和 S 矩阵理论的一种努力)等问题的原创性论文及其他有关科学史方法论、科学社会学及后现代思潮的论文 50 多篇，并编有《量子场论的概念基础》(*Conceptual Foundations of the Quantum Field Theory*，剑桥大学出版社，1999 年)和第 20 届世界哲学大会会议文集中的第 10 卷《科学哲学》(哲学文献出版社，2001 年)。

出版前对本书的评价

在这部深奥但可读性强的专著中，曹天予教授注意到，对微观世界所有观测现象的一个逻辑一致的描述即将出现。这个所谓的标准模型(standard model)是在实验和理论复杂的相互作用中涌现出来的。在最近几十年中，具有极大毅力和丰厚资助的实验学家识别并且研究了那些显现为物质的基元的东西及其受到的力。同时，理论学家创造和提炼了一种数学框架——量子场论(quantum field theory)——据此，标准模型得以表达。在所有与实验有关的方面，标准模型出现凯歌高奏的局面。但到目前为止，仍然有太多的问题悬而未决。很多理论物理学家相信，一种全新的理论体系——超弦理论(superstring theory)或其他类似的理论——是需要的。

在这些发展的进程中，当前理论的概念基础已经变得模糊。大多数物理学家过分关注他们探索的现象或建立的理论，以至于不再为其学科的逻辑基础和历史起源而费神。大多数探索科学知识和客观实在之关系的哲学家装备不佳，不足以处理量子场论中错综复杂

的事物，更不用说是它的后继理论。曹天予告诫我们，在能够建立一个更好的理论，抑或理解标准模型的真实意义之前，首先必须理解我们身处何处以及我们是如何到达这一步的。他对经典场论和量子场论的发展和诠释作出了清晰的论述——在所有自然力的规范场论(gauge field theory)的创立中，这一论述达到了登峰造极的地步——这将使物理学家和哲学家都关注科学是什么以及它是如何演化的问题。

——格拉肖(Sheldon Lee Glashow)，

哈佛大学

曹天予的著作面向当今物理学的一个基本问题：场论不再在基本运作层面上推进我们对大自然的理解。原则上，当前的场论根据粒子物理学的量子"标准模型"和引力的经典牛顿—爱因斯坦模型(Newton-Einstein model)解释了所有观测到的现象。但是，这种结构显然是不完善的：粒子物理学的成功是建立在特设的未作解释的假说基础之上，而且经典引力理论未被整合到非引力现象的量子描述之中。在一个世纪之中，场论提供了当今基本物理定律的框架，而在场论范围内推动这些问题的解决已经不可能了。在没有新的实验引导理论猜测的情况下，一些理论学家已经采用数学上优美和必要的概念来指导他们发明一种新的物理思想：弦论(string theory)。但是，人们还没有能力在实验上检验弦论是否能够成为量子场论的行得通的取代者，这就使得其他物理学家感到沮丧，而且促使科学期刊的编辑对我们的科学实在论承诺提出质疑。在场论发展的关键时刻，这本书对这一学科作了受人欢迎的综述，清楚地记载了场论发展到今天的历程以及相互冲突的观点。曹天予属于完全熟悉现代资料的技术复

杂性的那一类新物理学史和物理学哲学专家。他的讨论准确、详尽、富有洞察力。科学工作者和有兴趣的读者将在书中看到当代基础物理理论的一种精彩描述。科学史学家和科学哲学家将在书中发现对潜藏在思想背后的逻辑的深刻讨论。职业研究者和学生将会欣赏到对我们的学科历史的经过审慎研究后的复述。

——贾基夫(Roman Jackiw),

麻省理工学院

这一作品是富有感染力的,它主要是将对现代场论(包括广义相对论、量子场论与规范场论)的很多技术细节的理解,与哲学上捍卫理论物理学的合理性和客观性的精神追求联系起来,这种精神追求在曹天予支持结构实在论的立场中得到体现。这部著作对当代科学史和科学哲学中流行的后现代主义思潮是一服强效解毒剂。对于要对场论的现代发展的复杂历史作出连贯理解的物理学家,和要把他们的判断建立在对这些论题的技术细节的精确介绍上的科学史学家和科学哲学家,本书是同样适用的。

——雷德黑德(Michael Redhead),

剑桥大学

追随马赫(Ernst Mach)和默茨(John Merz)的步伐,曹天予对20世纪基础理论物理学的概念发展,以及它们的历史根源与局限,作出了精彩、深刻的阐述。这是对物理理论所依据的形而上学的、哲学的和技术的假设的深刻的批判性研究,这些物理理论在目前的实验探测远未达到的领域内如此令人印象深刻地、精确地描述了自然界,它们是狭义和广义相对论、量子场论,尤其是作为标准模型的基本成分

的相对论性量子规范理论。如同马赫一样，曹天予的探究表明了对历史敏感的哲学研究是如何揭秘物理理论的，而且他的著作无疑是有助于指示基础物理学的理论化的未来道路的有价值的路标。物理学家、科学史学家和科学哲学家会热切地阅读这部著作。

——施韦伯(Silvan S. Schweber)，

布兰代斯大学

引自精装版原著的评价

以20世纪场论作为科学如何演进的一个范例，该书对之进行了深入研究。这本书对于任何研究场论的专家或者场论成就的赞美者来说，都是必不可少的阅读材料。该书作为一次有水平的综合已广受赞誉。

——《欧洲核子研究中心快报》

这是一部具有非凡眼光的著作，它把物理学与科学史和科学哲学结合起来……有可能成为关于现代场论的历史和概念基础的标准参考书目。

——库欣(James T. Cushing)，

《今日物理》

这本书缀满宝石……是一部不寻常的著作……[它]把历史学知识与坚实的科学探究熔为一炉，聚二者之精华以阐明场论的核心概念及

其演变。该书在这一点上取得了巨大成功。在此基础上，曹天予提出了一种他称为“本体论综合”的科学变革的论述，这一论述组成了一种“结构实在论”的科学实在论。

——法因(Arthur Fine)，

《科学》

这部按照上述[几何、量子场和规范场]纲领来建构的著作，是对于20世纪场论的概念起源和发展的一个最新的、博识的和详细的历史说明……围绕这一丰富的历史内容的是具有哲学特质的导言和结尾部分。对我来说，它们是这本书最能激发灵感的部分。

——艾奇逊(Ian Aitchison)，

《自然》

曹天予写了一部非同寻常、富有教益的著作，对一些传统的观念提出了挑战。

——布朗(Laurie M. Brown)，

《美国物理学杂志》

他的工作是精确的，文献确凿，包含了对规范场论的曲折发展的详尽描述……这不仅对于物理学家，而且对于物理学史学家，同样都是一个具有非凡意义的研究……这是一部信息丰富、引人入胜的著作。正如《欧洲核子研究中心快报》(*CERN Courier*)的编辑弗拉泽(Gordon Fraser)告诉我的：“我的办公室少不了这本书!”

——韦尔特曼(Martinus Veltman)，

《物理世界》

对当代物理学的哲学理解作出最杰出贡献的著作之一……像曹天予的这样一部著作，可能有助于理论物理学学生与从事研究的青年科学家发现探索统一理论的新道路。

——多拉托(Mauro Dorato)，

《认识论》

曹天予关于20世纪场论的著作……抱负非凡，极其成功地尝试对物理学是如何在最基本的层面上发展起来给出连贯说明……曹天予的处理是见闻广博且造诣深厚的，然而对于大多数物理学学生和许多科学史学家与科学哲学家来说还是易于理解的。

——克拉格(Helge Kragh)，

《伊希斯》

一部开创性的划时代的物理学史杰作

非常高兴看到曹天予教授撰写的《20世纪场论的概念发展》中译本，在上海世纪出版集团科技教育出版社出版。

这是一部杰作。

这是自然科学史领域内一部开创性的而且是划时代的著作。正是此书完整、翔实、准确而且首先分析和总结了“20世纪场论的概念发展”；并根据大量历史发展的史实，在崭新的科学事实、科学理念的基础上，深入、具体地探讨了人类认识物理世界的历程和客观规律。过去人们所认识的物理世界，是由物质的“三态”(即固态、液态、气态)组成的宏观物理世界；后来就逐步深入、深化为分子、原子、原子核、质子、中子、电子、中微子等粒子的系列；现在却发现还要加上“场”的理念，由重力场、电磁场发展为胶子场、中间玻色子场，还有有待超大型对撞机确证的希格斯场……而极令人感兴趣的是，原来被认为属于“粒子”系列的分子、原子、原子核、质子、中子、电子、中微子等，竟然也“演化”为场的量子激发态的表现；因

此，物理世界，也就是物质世界，就此统一在一个新的“图景”，统一于由“场”的理念所描述、所刻画的物质世界!

毫无疑问，这是人们对“物质”世界认识的大转变，也可以认为是大变革。不连续的“粒子”世界，竟然转变为连续的“场”的世界；连续介质的“场”的物质观，竟然取代了不连续的“粒子”的物质观！所有这一观念上的“大转变”、“大变革”，究竟是怎样发生的？这就是本书所要回答的中心内容。

如果略为追踪一下认识史，在20世纪以前，人们多半把“场”的理念，看作是空幻的“猜想”，有些人更看作是“精神”领域中的“事物”。自20世纪以来，由于一个世纪来一连串大物理学家的共同努力，现在已完全确立了以所谓“规范场”为纲领的量子场论的“标准”模型！！！当然，还要回答：这一“标准”模型，将往何处去？将会碰到什么新问题？有哪些“理念”，需要进一步深化、发展，甚而要作哪些修改？！这就是本书所要探讨和回答的问题。

重要的是：本书并不是以某种“虚无缥缈”的“猜想”，或“虚构”某种“高论的胡说”，来探讨、回答这些问题，而是尽可能“立足”于现代科学基础，来分析和研究这些问题。这就是本书可贵之处!

本书包含下列一些内容。

一、首先是“经典场论的兴起”，尤其是古典的“场”的概念或“场论纲领”的确立

在本书的第57页，作者曾扼要地做如下的概述：“经典场论的兴起，在寻求表观超距作用的动力因(efficient cause)方面有很深的渊源。就电磁学来说，威廉·汤姆孙与麦克斯韦在他们的以太场论

中，通过引入新的实体(电磁场)与新的本体(连续以太)，成功地解释了超距作用。场具有能量，因而代表物理实在。但是，作为机械以太的一种状态，场不能独立存在。在洛伦兹的电动力学中，场仍然是以太的一种状态。但是，既然洛伦兹的以太被剥夺了所有物质属性，变成了虚空(void space)的同义词，那么场就享有独立的本体论地位，与物质平起平坐。因此，在物理学研究中出现了新的研究纲领，一种建立在场本体基础上的场论纲领，与建立在粒子本体(与空间和力一起)基础上的力学纲领形成对照。”

凡是在大学里学过麦克斯韦电磁定律的学生，对于这里的叙述，自然是很熟悉的！例如，早在60年前，我在清华大学物理系学习电磁学时，叶企孙老师就向我们说过：“场是以太的力学的某种状态*，还是真实的物理实在，在电磁学范围内，那是吵不清楚的！”曹天予教授在本书就更具体地回顾了这一历史进程。在本书的第15页，他具体写道：“在19世纪的物理学中，场的实体性是一个有争议的论题。有时人们论证道，麦克斯韦确立了场的实体性，因为他证明了能量存在于场当中。但是，这种论断是有问题的。对于麦克斯韦，场不是一种客体，而只是遵循牛顿运动定律的机械以太的一种状态。这意味着对于麦克斯韦来说，场不是自存的，因而不可能是物质的。能量在场里面的存在所确立的只是以太的实体性，而非场的实体性。”由于“机械以太的去除，使得场的实体性的去除成为必要”，“正是机械以太的去除，才建立了场的非物质的实体性。”而且，根据作者的意见，“连续的实体场是世界的基本本体的假设必须被视为场论的第一个基本信条，虽然在物理学史上并不

---

* 指以太的应力和应变。——何祚庥

总是这样的"。

接着，曹天予教授又写道："当电磁场被看作是世界的基本本体，而不只是一种数学装置或机械以太的一种状态时，存在着比这表面差别更深刻的东西。场是一种新型的实体……场因其非机械的行为而既不同于物质个体，也不同于机械以太。这种新的非机械本体的引入，开创了一个新的纲领——场纲领。场纲领不同于机械纲领之处在于：首先，它抛弃了机械以太，致使电磁相互作用的传输不能在机械纲领中得到解释。其次，它引入了一种独立的、不能还原为机械本体的电磁场实体。而这就为场纲领的进一步发展铺平了道路（见第2章）。"

我以为作者的这些论述是十分重要的。如果"场"不被看作是某种"世界的基本本体"的话，那还有什么对物质世界场的纲领的研究？！

但是，事情的发展，远没有"圈外"人士想象的那么简单，因为必须解释"电荷"的存在。例如，描述静电场势能 V 的方程通常会写成，$\nabla^2 V + 4\pi\rho = 0$，但势能 V 是和电荷密度 ρ 有联系的。如果我们把势能 V 看成描述"世界的基本本体"的某种方程式，那么就要回答电荷密度 ρ——其特殊情况会蜕化成一个点函数 $\delta(\xi)$——是什么！电荷的理念，能在"场"纲领里加以解释吗？在电磁场的"以太"模型中，电荷可看作以太的流体介质动力学中的"源"（source）和"汇"（sink），而如果"以太"一旦被取消，人们将如何解释这些"源"和"汇"？

在第50—53页上曹天予教授介绍了洛伦兹的工作。洛伦兹的以太"摈弃了所有机械的性质，因而完全从物质中分离出来。在这个框架中，电磁场被当作是以太的状态。既然以太没有机械性质，其

性质只是如同构成电磁场与物质之基础的虚空一样，那么电磁场就享有与物质一样的本体论地位。即，它代表着独立于物质的一种物理实在，而不是物质的一种状态，像物质一样具有能量，因而具备非机械实体的资格”。

但是洛伦兹并没有完全抛弃以太的概念。因为还需要解释“物体”，也就是指场体系中的“源”和“汇”，在运动时如何能保持内部的力的平衡。“洛伦兹假设，当物体在以太中运动时，除了电磁力，决定物体大小的分子力也受到影响……作为物理现象的原因，分子力的变化被设想为是由以太与物体分子之间一些未知的物理相互作用引起的，因此以太必定是实体的，这样它才能与分子相互作用。”

洛伦兹的“以太”也取得某些成功。因为他“据此试图对1887年迈克耳孙—莫雷实验(Michelson-Morley experiment)的否定结果作出清楚解释。”但是洛伦兹的理念，也遭到了许多批评，“庞加莱(Poincaré)批评洛伦兹……违背了牛顿第三运动定律”。

但是洛伦兹的杰出成就是：“洛伦兹的工作在电磁世界观对机械世界观的文化交锋中很有影响，并起了决定性作用。”而且，爱因斯坦在“晚年曾经对洛伦兹在发展场论过程中走的这一步的重要性作出评论：‘这是令人惊奇与大胆的一步，要是没有它，以后[场论纲领]的发展将是不可能的。’(Einstein，1949)”

为什么要在这一短“序”里，较长地引用了作者对“场”的观念发展的某些概述？原因是：1)这是“场纲领”或非机械的电磁场模型取代机械的或力学模型的“基础”，而这是物理观念的重大转变；2)这样的观念的转变——并不如通常人们所设想的那样——不能简单地归结为由于狭义相对论的出现，就此从物理世界力学模型转变为电磁场模型；3)但是，本书作者却将这一转变过程细致地勾画了出来，

这对于我们这些仅熟悉“近代”物理的发展，不熟悉“古典”物理的后学者来说，是一种教育和学习；4)顺便说说，本书对这类观念上大转变的考证还有很多很多，不能一一细述，现在只较详细地举出一例，而且是易于理解的一例，足证此书的价值。

二、 走向广义相对论和物理学研究的“几何纲领”

狭义相对论获得了空前成功；进一步发展是走向广义相对论。随着广义相对论的三大观测的验证的出现，广义相对论也就成为物理学界所认同的理论。但是，广义相对论是有重大争议的理论，尤其是广义相对论的进一步发展，爱因斯坦尽毕生之力追求的统一场论*，却没有成功！那么，人们将怎样看待广义相对论，怎样看待伴随广义相对论的成功而出现的“场”的几何化纲领？本书对这些问题也有相当细致而精辟的叙述。

广义相对论的基本理念，是用时空扭曲解释引力。广义相对论所以获得成功，主要是基于下列实验事实：惯性质量和引力质量存在着严格的比例关系。爱因斯坦认为，“一个物理系统的惯性质量与能量，似乎表现为同一个东西”，“惯性质量与引力质量之间的比例性，在目前所能达到的精度内，对于所有物体都是普适有效的；……我们必须把它作为普适有效的来接受它。(Einstein, 1907b)”正是在上述思想指引之下，爱因斯坦作出如下假定：“从现在开始，我们同意假设，引力场与参考系里的加速度在物理上是完全等价的(1907b)。”“爱因斯坦在论文(1912b)中把这个陈述称为等效原理(简称 EP)，在论文(1919)中，把这个思想称为‘我一生中最愉快的

* 实际上只限于引力和电磁力的统一。——何祚庥

思想’。”进一步，爱因斯坦把引力场归结为时空的弯曲本质。到了统一场论，爱因斯坦则进一步也希望将电磁场归结为时空的某种特性。这就是“场”的“几何纲领”的由来！

问题是：电磁场和引力场之间并不存在像“惯性质量和引力质量”那样的“同一性”，想从时空性质来统一解释引力场和电磁场，只能认为是缺乏实验基础的一种假想。历史上有不少物理学大师作出种种尝试——在《20世纪场论的概念发展》一书有较详细的介绍——但均没有获得成功。虽然爱因斯坦及其追随者在物理学基本理论方面作出许多划时代的贡献，但总的来说，场体系的“几何化”是一个失败的纲领。

极有意义的是：在“场”的“几何纲领”方面，曹天予教授却注意到“几何纲领”存在强版本和弱版本的区别。

“根据强版本，(i)时空的几何结构本身是物理上实在的，如同物质与电磁场的实在性一样，而且具有作用于物质的实在效应；(ii)引力相互作用被看作物质运动的时空几何效应的局域测量，作为时空曲率的一种表现，并通过测地线偏差方程得到表达。”至于“几何纲领的弱版本拒绝时空结构的独立存在，只是把它们看作场的结构性质。爱因斯坦在他追求统一场论的晚年，坚持这种立场。但是，从逻辑上讲，弱版本不必预设统一场论。它所预设的是时空的几何结构在本体论上是由物理(引力)场构成的。”

为什么在“几何纲领”中会存在这两种版本？原因还在于“爱因斯坦为等效原理列出的唯一经验证据是，‘所有物体在引力场中经历相同的加速度’。这个事实所支撑的只是所谓的弱等效原理(weak EP)，弱等效原理要求力学实验在一个匀加速系统中与在一个处于匹配的均匀的引力场里的静止系统中遵循相同的过程。但爱因

斯坦主张的却是所谓的强等效原理(strong EP)，强等效原理假设两个这样的系统是完全等价的，以至于不可能借助于任何一种物理实验来区分它们。”

对于等效原理这两种不同的理解，导致场体系几何化也出现“强”和“弱”两种不同的纲领。

为什么这一“强”和“弱”的区分如此重要？因为这涉及如何理解哲学实在论。实在论的基本思想是：世界统一于不随人们的主观意志而转移的客观实在，而时间空间却是客观实在存在的形式。重要的是，*形式并不等同于内容*。但是，在爱因斯坦的“强”几何纲领中，却把“时空的几何结构≡物理实在”。对此，曹天予教授历述了爱因斯坦在“引力理论的几何化”方面所获得的成就，但同时又指出爱因斯坦的缺失——“将时空几何引力化”的缺失——并公正地指出，这是爱因斯坦走向统一场论、但走向失败的根本原因。

其实爱因斯坦并非一点也没有察觉“时空几何引力化”或“时空几何的场化”的假想里存在着某些原则性的困难。1915年，著名的广义相对论学者施瓦西曾给出“真空中一个静态球对称质点的近似引力场解”，到了1916年，又得到了“一个孤立质点的*精确解*”。这些“近似解”和“*精确解*”均具有一个特征：当球坐标 R 趋向无穷大时，亦即在无限远处的度规，一定是闵可夫斯基度规。这意味着广义相对论并没有能完全做到“时空的引力化”。至少，在 $R \to \infty$ 大时，时空是空无所有或没有任何场的时空，亦即“时空是客观实在存在的形式”。

所以，在现代场论学者看来，广义相对论只是描述引力场的理论，原因就在于厄特沃什(Eötvös)有关引力质量和惯性质量成正比的实验，所证明的仅仅是“弱”等效原理。

三、 物理学研究中的量子场论纲领和量子规范场论纲领的确立

本书的第二篇和第三篇都是讨论长期以来物理学所追求的一个宏伟目标：物理世界能否在量子化的场的理念里，统一地由“场”来描述，而且是以某种“统一的场”来描述？

其实，在上述“走向场的几何纲领”的追求中，即使已“做成”引力场和电磁场的“统一”场论，但场的“源”和“汇”的问题，其实仍需要在“场”纲领以外分别加以处理。洛伦兹曾经有如下尝试：用电磁场来解释电子质量，而且还能解释洛伦兹收缩！但是，即令电子质量可以用电磁场的“凝聚”来解释，但不能由此而推广到质子、中子、μ 子、中微子等其他粒子！极为有意义的是，这一“粒子”观和“场”观的矛盾，竟然在量子化的“场”的理念中得到较完满的解决——但同时又提出许多原则性问题！本书对这一历史过程做了详细的追踪和回答。

本书的第二篇和第三篇所涉及的内容有：a)量子论的建立，即量子力学及其核心问题，波粒二象性理念的建立；b)量子场论的基本概念的形成，其中包括玻色子场的量子化和更为困难一些的通常称为实物的费米子场的量子化；c)可重正量子场论的形成；d)规范不变性和规范场概念的确立；e)走向量子规范场论；f)走向标准模型和量子规范场纲领的确立——这实际上是几何纲领和量子场纲领的综合，因为从几何学的观点来看，规范场是和“纤维丛”相联系的一种几何学。

对于我们这些长年从事粒子物理学和量子场论的研究工作者来说，对这段长达五六十年的现代物理学发展史，无疑既饶有兴趣又耳熟能详。但是，本书的突出优点在于：不仅促使我们“温故”，而且还启示我们“知新”。下面仅举两例：

事例一，是此书对韦尔特曼工作的评价。量子规范场论有两大

难点：一是如何走向量子化，这首先是由法捷耶夫和波波夫用“费恩曼路径积分+法捷耶夫—波波夫鬼态”所回答和解决的；二是如何进行重正化，而人们所熟知的“事实”是由特霍夫特发展的维数正规化方法所回答和解决的。实际上更准确的历史是，首先是韦尔特曼提出规范场论如何进行重正化的许多理念，并最终由特霍夫特完整地实现。所以，在此书出版后的1999年，韦尔特曼和他的学生特霍夫特共同获得了诺贝尔物理学奖。

事例二，标准模型，亦即弱电统一模型的建立，有赖希格斯引进的希格斯机制和希格斯粒子。至于希格斯工作的先驱——南部阳一郎所做的真空自发破缺的经典工作——却为人们所忽略。但是，此书却公正地作了叙述，准确地评价了南部阳一郎所做的历史性贡献。2008年，南部阳一郎也获得了诺贝尔物理学奖。

令人惊讶的是，在本书的第二篇和第三篇中，曹天予教授不仅准确无误地叙述了量子场论、量子规范场论的基本概念以及这些概念的发展和变化，而且对各个阶段不同学者的贡献均给予了符合历史事实并且是科学的评价。这是很不容易做到的事情，曹天予教授竟然做到而且极好地做到了！

所以，此书得到一大群著名物理学家(其中有些是诺贝尔物理学奖获得者)的极高评价。下面是著名诺贝尔物理学奖获得者格拉肖教授和威特曼教授对此书的评述：

> 在这部深奥但可读性强的专著中，曹天予教授注意到，对微观世界所有观测现象的一个逻辑一致的描述即将出现。这个所谓的标准模型(standard model)是在实验和理论复杂的相互作用中涌现出来的。在最近几十年中，具有极大毅力和丰厚资助的实验学

家识别并且研究了那些显现为物质的基元的东西以及其受到的力。同时,理论学家创造和提炼了一种数学框架——量子场论(quantum field theory)——据此,标准模型得以表达。在所有与实验有关的方面,标准模型出现凯歌高奏的局面。但到目前为止,仍然有太多的问题悬而未决。很多理论物理学家相信,一种全新的理论体系——超弦理论(superstring theory)或其他类似的理论——是需要的。

在这些发展的进程中,当前理论的概念基础已经变得模糊。大多数物理学家过分关注他们探索的现象或建立的理论,以至于不再为其学科的逻辑基础和历史起源去费神。大多数探索科学知识和客观实在之关系的哲学家装备不佳,不足以处理量子场论中错综复杂的事物,更不用说是它的后继理论。曹天予告诫我们,在能够建立一个更好的理论,抑或理解标准模型的真实意义之前,首先必须理解我们身处何处以及我们是如何到达这一步的。他对经典场论和量子场论的发展和诠释作出了清晰的论述,在所有自然力的规范场论(gauge field theory)的创立中这一论述达到了登峰造极的地步,这将使物理学家和哲学家都关注科学是什么以及它是如何演化的问题。

——格拉肖(Sheldon Lee Glashow),

哈佛大学

他的工作是精确的,文献确凿,包含了对规范场论的曲折发展的一个详尽描述……这不仅对于物理学家,而且对于物理学史学家,同样都是一个具有非凡意义的研究……这是一部信息丰富、引人入胜的著作。正如《欧洲核子研究中心快报》的编辑弗拉泽

(Gordon Fraser)告诉我的："我的办公室少不了这本书!"

——韦尔特曼(Martinus Veltman)，

《物理世界》

此外，还有一大批我们很熟悉的量子场论专家和量子场论专著的作者对此书的评述：

在场论发展的关键时刻，这本书对这一学科作了受人欢迎的综述，清楚地记载了场论发展到今天的历程以及相互冲突的观点。曹天予属于完全熟悉现代资料的技术复杂性的那一类新物理学史和物理学哲学专家。他的讨论准确、详尽、富有洞察力。

——贾基夫(Roman Jackiw)，

麻省理工学院

曹天予对20世纪基础理论物理学的概念发展，以及它们的历史根源与局限，作出了精彩、深刻的阐述……如同马赫一样，曹天予的探究表明了对历史敏感的哲学研究是如何揭秘物理理论的，而且他的著作无疑是有助于指示基础物理学的理论化的未来道路的有价值的路标。

——施韦伯(Sivan S. Schweber)，

布兰代斯大学

这本书缀满宝石……是一部不寻常的著作……[它]把历史学知识与坚实的科学探究熔为一炉，聚二者之精华以阐明场论的核心概念及其演变。该书在这一点上取得了巨大成功。在此基础

上，曹天予提出了一种他称为“本体论综合”的科学变革的论述，这一论述组成了一种“结构实在论”的科学实在论。

——法因(Arthur Fine)，

《科学》

这是一部具有非凡眼光的著作，它把物理学与科学史和科学哲学结合起来……有可能成为关于现代场论的历史和概念基础的标准参考书目。

——库欣(James T. Cushing)，

《今日物理》

这本书对于任何研究场论的专家或者对场论成就的赞美者来说，都是必不可少的阅读材料。该书作为一次有水平的综合已广受赞誉。

——《欧洲核子研究中心快报》

需要指出的是：这些赞誉或者出自著名理论物理学者，或者出自极有影响的物理刊物的书评，而且用了许多高级的形容词。然而，这绝非虚誉！应该说，本书的精彩内容太多太多，无法一一备述，所以只好摘录一些已获得的赞誉作为此书的介绍。

四、本书的重大贡献——以大量的物理学史实，确立了结构实在论

本书除了对物理学、物理学史作了巨大贡献以外，还对科学哲学也作了重大贡献——以大量的物理学发展史实，科学地确立了结构实

在论。用中国人喜欢用的哲学语言来说，本书丰富和发展了唯物主义哲学，自然界的实体或物理实在已统一于有“规范”不变特征的“量子”化的“场”。只不过在西方哲学著作里，他们认为唯物论是一种低级的、庸俗的、品格不高的哲学名词。

有一个值得讨论的问题是：本书有没有缺点？

本书是一部开拓性的著作。任何开拓性或首倡性的科学工作，包括科学史，总是有缺点的。我以为本书在总体上是为了阐明作者所持的自然观——结构实在论——而如果自然界的某种物理“结构”构成了“实在”的话，那一部“20世纪场论的概念发展史”，就不能仅限于讨论场的“理论”框架的演变和确立，而是还要看到构成“场”的“实体”，是如何从纷乱的实验事实里涌现出来的。没有电子、光子就不会有量子电动力学，没有夸克和胶子和它们所满足的“色”规范群，就不会有量子色动力学。没有中间玻色子，没有希格斯机制及其相应的希格斯粒子，就没有弱电统一理论；甚而这一理论的可重正化性问题，也有赖于夸克和轻子的对称性，才能消除三角形曲线图的反常。

2005年，已故著名物理学家、“两弹一星”元勋、我们这代许多理论物理学工作者的老师彭桓武教授，在谈及科学研究与创新时曾意味深长地引述了两段爱因斯坦名言：“Pure logical thinking cannot yield us any knowledge of the empirical world. All knowledge of reality starts from experience and ends in it .” “We now realize, with special clarity, how much in error are those theorists who believe theory comes inductively from experience. Even the great Newton could not free himself from this erroe (Hypotheses nonfingo).”

有不少人翻译过爱因斯坦的这两句话。我所喜欢的译文是：

“纯粹的逻辑思维不能给我们关于经验世界的任何知识；一切关于实在的知识，都是以经验开始，以经验结束。”“我们现在特别清楚地认识到，那些相信理论是从经验中归纳出来的理论家是多么荒谬，甚至伟大的牛顿也不能从这种错误中摆脱出来！伟大的牛顿就说过，‘我不作假设’。”

据此，彭桓武教授指出：“爱因斯坦的两段话是互补的，前段话强调理论对实验的依赖性，后面这一段强调单纯从经验进行归纳是不够的，从经验到理论的飞跃很多都是从猜想和假说中建立起来的，‘不作假设’是不对的……我在北大曾经讲过，在北戴河《物理通报》的会上也讲过，我在中学时曾问老师凸透镜公式 $\frac{1}{f}=(n-1)\left(\frac{1}{R_1}+\frac{1}{R_2}\right)$ 是怎么从经验中归纳出来。当时老师也答不上来，用归纳法得不出来嘛！后来发现这公式是从理论推导出来的。我过去一直是相信理论从实验归纳出来的，其实上述凸透镜公式是由实验所确定的折射定律，加上逻辑推演的或数学的推导才得出的！”

彭桓武教授还指出：“在中国比较缺少逻辑思维。你看《孟子》中什么人性是向善的，水是向下的。在前面没有个大前提，二者有何联系？人性不是物质；水可以向下流，一个球也可以向下滚；能说人性有多少重量？……我在英国留学时遇到一位爱丁堡大学的数学教授惠特克(Whittaker)，他就说‘中国缺少逻辑思维！’”

我以为，曹天予教授所总结的“20世纪场论的概念发展史”，以大量的史实充分证明爱因斯坦所持高超的见解，是何等切实而重要。当然，量子规范场论的确立并不是认识的终点，人们对自然界的实体及其运动、变化、发展规律的认识，还在继续前进。但是，正如我们的另一位前辈理论物理学家，也是我们这代理论物理学工作者(包

括彭桓武老师在内)的共同的老师周培源教授在1991年《我的人生道路与治学观》一文中指出："对于任何问题，譬如在对待科研工作中的论据和论点的科学性方面，我都十分注意实事求是问题，不能浮夸，不能掺入半点虚假。为此，我曾提出一个新的科学理论所必须同时满足三个条件：一要能够说明旧的科学理论所能够说明的科学现象；二要能够解释旧的科学理论所不能解释的科学现象；三要能够预见到新的科学现象并能够用科学实验证明它。如果某个理论或结论有不实之处，或人为地掺入了非科学的成分，就不可能同时满足这三个条件，特别是第三个条件。我想以此来检验我的科研工作结果的'纯度'。"

愿以周培源教授的"名言"，和作者、读者共勉！

何祚庥

2008年12月

一部为科学实在论作辩护的当代物理学思想史

曹天予，20世纪60年代初在北京大学哲学系肄业。以后，在坎坷苦难的历程中，他自学了理论物理学直到量子场论。1983年，他到英国剑桥大学留学，在科学哲学家赫西(Hesse)和雷德黑德的指导下攻读科学哲学和科学史。在1987年获博士学位后到美国从事研究。现在美国波士顿大学哲学系任教。

他积十余年的辛勤研究，写出了《20世纪场论的概念发展》这部著作，于1997年在英国剑桥大学出版社出版。在此书出版前，评审者就对本书作出了很高的评价。哈佛大学的物理学家格拉肖认为，他关于经典场论、量子场论、规范场论的清晰阐述“将使物理学家和哲学家都关注科学是什么以及它是如何演化的问题”。麻省理工学院的贾基夫认为，在场论的发展的关键性时刻，“这本书对这一学科作了受人欢迎的综述，清楚地记载了场论发展到今天的历程以及相互冲突的观点……他的讨论准确、详尽、富有洞察力”。雷德黑德认为，这部著作“为理论物理学的合理性和客观性作了生动有力的

哲学辩护”，“对当代科学史和科学哲学中流行的后现代主义思潮是一服强效解毒剂”。布兰代斯大学的物理学家施韦伯认为，“追随马赫和默茨*的步伐，曹天予对20世纪基础理论物理学的概念发展，以及它们的历史根源与局限，作出了精彩、深刻的阐述……是有助于指示基础物理学的理论化的未来道路的有价值的路标。”

此书出版以后，美国《今日物理》(*Physics Today*)、《科学》(*Science*)、英国《自然》(*Nature*)、《美国物理学杂志》(*American Journal of Physics*)、英国《物理世界》(*Physics World*)等著名刊物分别发表了物理学家库欣、哲学家法因、物理学家艾奇逊、布朗、韦尔特曼撰写的书评。欧洲核子研究中心(CERN)的《欧洲核子研究中心快报》发表了书讯。该刊的主编弗拉泽说，“我的办公室少不了这本书。”据说，日本、德国、俄国的刊物也已发表了书评，法国也即将发表书评。一部学术著作，出版不久，就在国际学术界引起如此广泛的反应，这是不太常见的吧。

在当今的西方世界，反实在论思潮十分盛行。正如曹天予在前言中所说，反实在论有长久的传统，其中包括经典的实证论、建构经验论和工具论等等。库恩的不可通约性命题否认了任何理论本体论的实在性。科学社会学的强纲领认为科学是一种社会建构。后现代主义的、激进的社会建构论认为科学只是一种说服、处理和制造事实与知识的修辞艺术，知识是一种权力举措和产物，与真理和客观性无关。普特南的“内在实在论”拒绝有关理论实体的客观实在性的外部问题，实质上也是一种反实在论。曹认为，对于物理学来说，有关自然界的基本组成和行为的实在论假设是必不可少的。在运用假

* 《19世纪欧洲思想史》的作者。——范岱年

说-演绎法的理论物理学中，如果没有力、场、以太、刚性或动力学时空、虚量子、夸克禁闭、规范势这些曾被称为形而上学的假设或假想实体，就无法前进一步。也许正是曹的这种反潮流精神，引起了国际学术界特别是物理学界对这部著作的广泛兴趣。

曹在第1章"导论"中承认，科学作为文化的一种形式，不能与社会相分离。科学所用的语言是社会交往的产物。科学所研究的问题、动机、材料、手段都由社会所提供。科学的机构和体制得到社会经济结构的支持和制约。而科学家的思想、比喻、概念方案也都来自他们的文化环境。因此，科学知识不能不受社会、文化、历史的影响和制约。但是，社会建构论者在强调科学是社会的建构物时，却忽视了"科学活动的目的是认识自然"这一关键之点。物理学的历史表明，大多数大物理学家从事研究的目的就是探索真实世界的图景。曹在本书中对20世纪各种场论作批判性的历史考察，就是要探测这些理论中物理学家的世界图景演变的模式和方向，揭示在这些理论中的本体论转变和本体论综合。

曹把这部概念史分作三篇，即关于基本相互作用的几何纲领、量子场纲领和规范场纲领。在这三篇之前的第2章中，他讨论了"经典场论的兴起"，指出：从法拉第到麦克斯韦，他们的电磁场理论仍是力学的以太理论。他们只把物质和以太看作具有本体论地位的实体，不把电磁场看作实体，电磁场只是以太的一种力学状态。直到洛伦兹，才把电磁场和物质看作独立的物理实在，而以太是一个空虚的绝对空间参考系。正如爱因斯坦所说，"这是令人惊奇与大胆的一步，要是没有它，以后[场论纲领]的发展将是不可能的。"

在以广义相对论为中心的第一篇"基本相互作用的几何纲领"中，曹首先指出，爱因斯坦在狭义相对论中摒弃了洛伦兹的以太和绝

对空间。在广义相对论中，最先，爱因斯坦追随马赫，把可称重物体看作唯一的物理实在，它们决定了引力场和时空的几何结构。由于引力场方程的真空解的发现，这种诠释站不住脚了。以后，外尔和爱丁顿把时空的几何结构看作物理实在，而把引力归结为时空流形曲率的表现。曹把这种观点称为强几何纲领。但爱因斯坦从未接受这种观点。后来，他在试图把引力与电磁场相统一的统一场论中，把总场的一部分引力场看成是最终的物理实在，而时空是它的结构性。曹把这称为弱几何纲领。所以，爱因斯坦并不是把引力几何化，而是把时空几何引力化。曹还介绍了几何纲领的进一步发展。例如，外尔在他的统一场论中，把电磁势守恒与时空标度的规范不变性相联系(1922)，以后又改为把它与电磁中量子相位变换的定域规范不变性相联系(1929)，这是以后规范场理论的先驱。又如，卡卢察和克莱因在他们的统一场论中，则把电磁势当作五维时空度规张量的部分分量。近年来，人们在试图建立把引力与其他三种相互作用实现统一的超统一理论时，又重新对卡卢察的高维空间统一场论发生了兴趣。对于几何纲领这一部分，韦尔特曼和法因在他们的书评中都作了很好的评价。

在第二篇“基本相互作用的量子场纲领”中，曹首先介绍了量子论的兴起，并指出量子论似乎是削弱了场纲领的基础。因为：①量子论给能量在空间中的连续分布一个极限，这是和场本体论相冲突的；②它违反了可分性原理；③量子论不允许粒子在量子跃迁时，或在它们的产生和湮没之间，有连续的时空路径，这和场纲领中传递相互作用的方式相冲突。而且量子论的概率诠释假设了粒子本体论。在多体问题两种量子化程序(二次量子化和场量子化)中，分别预设了粒子本体和场本体，是粒子和场两种本体的并存。1928 年，约当和

维格纳把描述单个费米子的波函数看作费米场，并实现其量子化，这样就让场本体取代了粒子本体，物质粒子(费米子)不再被看作永久的独立存在，而是费米场的瞬时激发，是场的量子。这样，一种新的场纲领——量子场纲领就开始形成了。

量子场论的场本体论又一次改变了关于真空的观念，真空像是充满了负能量电子的海洋，像是一种可激化的介质，真空是激烈活动(涨落)的舞台。量子场论也改变了传递相互作用的方式。相互作用不再由一种连续的场来传递，而由这种场的激发(离散的虚粒子)来传递。这些虚粒子与真实粒子局域耦合，并在真实粒子间传播。

由于不确定性关系，局域激发要求任意大的动量。于是，相互作用不是由单个虚动量量子来传递，而是由无穷多个合适的虚量子的叠加所传递。这无穷多个具有任意高动量的虚量子就带来著名的发散困难。而这是通过重正化来解决的。曹在“重正化”一节(7.6节)中，对各种发散的根源、重正化纲领的建立、重正化的理由作了详尽、深刻的论述。他指出：重正化的本质是把无穷大量纳入质量、电荷等理论参量中，这等于模糊了作为局域激发概念基础的严格的点模型。

在20世纪40年代，把量子场论应用于电磁场(也就是量子电动力学)，取得了惊人的成功。但把同一种方法应用到弱和强相互作用上时，却失败了，这样，在50年代到60年代，S矩阵理论(或色散关系理论)有了巨大的发展。这是一种唯象的分析手段。曹在“S矩阵理论”一节(8.5节)的末尾，总结了量子场论和S矩阵理论在基本假设方面的七大差别，并指出，自S矩阵理论产生以后，它与量子场论一直相互施加压力，理解这两种研究纲领之间的相互作用，对理解粒子物理学的概念发展，是很有帮助的。接着他又介绍了部分守恒轴

矢流假说和流代数，这是一种对称性方法，它首先应用在强子的电磁和弱相互作用领域，以后推广到低能强相互作用领域。

到了20世纪60年代，S矩阵理论又逐渐衰退。特别是由于60年代末、70年代初在理论和实验方面出现了有利于规范场论的巨大突破，物理学家的兴趣又转向了规范场研究纲领。

第三篇讨论基本相互作用的规范场纲领。曹首先介绍了规范不变性和确定相互作用形式的规范原理；并指出现代规范理论是从杨振宁和米尔斯提出的关于强相互作用的同位旋规范不变性理论开始的。但是由于无法得到说明核力短程性的规范不变机制，物理学家就放弃了这一理论。而更大的困难是如何使规范场量子具有质量。从50年代末到1964年，物理学家引入了自发对称性破缺机制，解决了这一困难。说明核力短程行为的规范不变性机制也由"渐近自由"和"色禁闭"加以解决。特别是由于1969年斯坦福直线加速器中心(SLAC)通过深度非弹性散射实验，发现了标度无关性定律，人们开始接受夸克—部分子模型和场论框架中的量子色动力学作为强子物理学研究的基本框架。在弱、电相互作用领域，随着1971年韦尔特曼和特霍夫特证明非阿贝尔规范理论可以重正化，粒子物理学家就设想建立一个弱电相互作用统一理论的可能性，该理论预言的中性流的存在也于1973年得到了实验证实。这样就大大增强了粒子物理学家对规范场论的信心。

在规范场纲领的指引下，物理学家提出了标准模型，它的内容包括强子的夸克模型，夸克与轻子的对应关系，以及用弱电统一理论和色动力学描述的动力学。标准模型在20世纪70年代末、80年代初取得了巨大的成功，发现了新的夸克和轻子；发现了传递弱相互作用的中间玻色子 W^+、W^- 和 Z^0，进一步确认了弱电统一理论；找到了

传递强相互作用的胶子存在的证据。但是自 80 年代中期以来，规范场纲领在进一步推广标准模型方面遇到了一系列困难。例如，对于 π 介子—核子低能相互作用，色动力学仍无法作出解释；弱电统一理论的自洽性仍有可疑之处；希格斯粒子是否存在尚无法确认。弱、电和强相互作用统一的大统一场论远未成功，至于引力场的量子化以及与其他三种基本作用相统一的超统一场论更看不出成功的希望，规范场纲领又处于停滞不前的境地。于是，钟摆又摆向另一个方面。在新方向方面，曹主要讨论了有效场论，而对新的超弦理论和超对称理论，则没有涉及。

曹天予在该书最后一章“本体论综合与科学实在论”中，结合 20 世纪场概念发展，得出了他对科学发展模式的观点。曹首先反对的是以库恩、费耶阿本德为代表的观点，因为他们否认科学发展中的连续性和进步，否认科学的发展愈来愈趋近于真理，因为他们认为科学革命前后的理论是不可通约的，两者用的是不同语言，描述了不同世界。曹认为他们的观点带有明显的相对主义和反实在论的含义。曹也不同意赫西、塞拉斯(Sellars)和劳丹(Laudan)等人对库恩观点的修正。赫西等人认为，在科学革命前后的相继理论之间，描述局部的经验定律和具体事实的语言还是可通约的，这些经验定律和事实描述在革命后的理论中仍可留存。在这方面，科学的发展表现出某种进步和积累。但他们否认科学理论中的普适定律和本体论在革命后能够留存并有所进步。曹因此称他们的观点是弱实在论。曹天予采纳并修正普特南的因果性指称理论作为自己反对不可通约性命题的理论框架，认为理论的“结构性”或“理论实体间的关系”虽然不能被直接感知，却是可通约的，而且对确定理论项的指称和其意义，具有关键性的作用。在革命前后的相继理论之间，通过本体论的综合，不

仅是经验定律和具体事实，而且理论的本体论的结构(例如内、外部对称性，可几何化性，可量子化性等)也可通约和继承。曹进一步指出："大多数科学家相信经验定律只有'局部'有效性，而本体论赋予科学以统一的力量。本体论作为世界基本结构的模型，是作为经验定律的基础的一般机制的载体。由于在它的基础之上，可以建立一门科学的统一概念框架，所以本体论在理论上比个别经验定律更为基本得多。"

曹认为，20世纪场论发展的历史表明，它们是通过本体论的综合来实现概念革命的。他指出：20世纪各种场论都渊源于经典电动力学，它是一种实体电磁场论，并以洛伦兹群为其对称群。引力场论是经典电动力学的直接后裔。爱因斯坦在创建广义相对论时，把洛伦兹不变性推广为广义协变性，通过等效原理把引力场与时空的几何结构相联系，从而开创了几何纲领。由此可以看出几何纲领与经典电动力学的本体论连续性。量子场论是经典场论的另一个直接后裔。其中经典电磁场被量子化电磁场所取代，又通过与玻色场的类比，引入了实体化的费米场。而基本的本体仍是某种实体的场。虽然量子场与经典场在结构性方面有根本性差异，但仍有一些共同性(例如，场的不同部分的可叠加性，个体化的不可能性)可作为场论的理论硬核。因而，在量子场论和经典场论之间，可以建立一种本体论的对应关系，从而看出两者的本体论之间的指称连续性。

规范场论是量子电动力学的直接后裔。杨振宁和米尔斯正是用与强相互作用场相关的局域SU(2)同位旋对称性来取代量子电动力学中的局域U(1)相位对称性，从而创建规范场论的，其中相互作用的形式均由规范不变性所确定。规范场和物质场一样均可量子化。所以，在规范场纲领和量子场纲领之间有明显的本体论连续性。另一

方面，规范场纲领和几何纲领之间的本体论连续性，则在于量子规范场和经典几何场在结构性方面也有指称连续性。这是因为几何纲领、广义相对论或其推广与变形，原则上可以取量子化的形式。* 而规范场可给予某种几何诠释。这样，两者就可以共有一些结构性甚至本质特征。在现代卡卢察—克莱因理论中，规范场纲领和几何纲领的本体论连续性就更明显了。

同时，曹天予也不同意以内格尔（Nagel）和波斯特（Post）为代表的科学发展观和收敛实在论。他们认为科学的发展是累积的、连续的，是单线式的进步，最后收敛于唯一的真理。曹认为他们的科学发展模式太简单化、太抽象了，不能说明科学发展的复杂情况。而曹天予所主张的本体论综合，既有继承和连续，也有变换和否定。不是把过去的本体论加以拼凑，而是有选择地、有变换地加以组合。例如，几何纲领是把等效原理和广义协变性思想相综合，而广义协变性虽来源于洛伦兹不变性但又不同于它。量子场论是经典场论和量子原理的综合，可是场的量子化思想虽来源于量子化原理但在本体论上不同于原子运动的量子化思想。规范场纲领是量子场纲领和规范原理的综合，可是规范原理虽起源于、但又不同于最小电磁耦合。在这三个案例中，以前的原理都是在变换成新的形式之后才被应用于本体论的综合。本体论综合时常把原来的原始实体变换成衍生实体或副现象（epiphenomenon），从而导致基本本体论的变革。例如，几何纲领把牛顿引力势看作度规场的表现；量子场论把经典场看作量子场的衍生物，而规范场纲领中的标准模型把电子电磁场看作量子化规

---

* 但实际上引力场的量子化还存在极大的困难。因为狭义相对论与电、弱、强三种相互作用结合成功的基础是这三种相互作用的耦合常数都是无量纲量，而引力相互作用的耦合常量是带有“质量”量纲的。我国规范场论专家汪容认为，“引力场”的不可量子化，已经成为20世纪物理学的一朵乌云。——范岱年

范场的衍生物。因此，本体论综合是一种实现概念革命的方式，而不是如内格尔和波斯特所设想的那样把革命前纲领中的旧本体论并入革命后新纲领的新本体论。另一方面，旧本体论所发现的世界的结构关系如内、外部对称性，可几何化性和可量子化性，在革命后的新本体论中仍留存着。概念革命通过变换与综合而实现，而不是绝对的否定。这是黑格尔所说的“扬弃”，有继承而又有否定。因为没有理由假设世界的结构性只能分为有限的层次，而且没有理论可以概括无限多的结构性，所以，一个终极的真的本体论概念是没有意义的。人们只能说“更好地”或“更接近于”真的本体论。

曹天予认为他的本体论综合概念胜过内格尔和波斯特的收敛实在论和赫西等人只承认“工具性”进步的弱实在论。因为本体论综合概念认为人们发现的世界结构是以辩证的方式继承和积累的，它更有力地说明了概念革命的机制和科学进步的模式。事实上，科学中的继承、积累和统一很少是以直接简单的形式实现的，而时常是通过变换与综合的方式实现的。通过变换的本体论综合体现了科学增长中变革与继承的协调，最适合于掌握科学史的不连续的外观后面的本质连续性，从而也可为科学实在论提供更强的论据。

曹的科学实在论就是结构实在论。这种实在论认为，基本实体的结构性和关系是本体论的一部分，而且是本体论中可以通过科学研究(通过把结构论断与假想实体关联起来的因果链)加以接近的唯一部分。虽然当理论变革并发现新的结构性时，这些结构性论断将被修正。但是这些结构性如同可观测性一样，是大体稳定并可累积的，因为，由于它们的可认识的同一性，在理论与理论之间是可翻译的。把外部对称性(如洛伦兹对称性)、内部对称性(如同位旋对称性)、可几何化性(与时空流形结构特征同形的结构性)、可量子化性(连续场

的结构性)等这类结构性用变换群及其不变性来描述，这是在19世纪末首先由庞加莱提倡并实现的，后来已成为爱因斯坦、狄拉克、维格纳、杨振宁、盖尔曼这些大数理学家的集体意识。而结构实在论正是这种集体意识的体现。

在坚持科学实在论的同时，曹天予也坚持了科学的合理性。他认为通过本体论综合所实现的科学革命，不是如同库恩所说的类似于宗教皈依的范式的格式塔转换，而是理性的有选择的组合。科学实在论认为，科学的目的是要对经验定律和世界的结构性作愈来愈真的论断，实现科学的进步。科学的活动是一种理性的活动。但是科学的增长和进步并不是单线式地逼近唯一的终极真理，而是辩证的综合。这种综合的观点不认为已有的成功理论必定是未来发展的模型。它要求科学家以开放的心态对待各种可能性。按照这种精神，场论的未来发展也不一定唯一地来自规范场研究纲领，也有可能要从别的根本不同的研究纲领(例如S矩阵理论)中吸取思想和技巧。

曹天予在前言中说，他的这部著作的对象首先是对理论物理学基础问题和历史透视感兴趣的学生，其次，是对当代自然观感兴趣的、受过基础科学教育的一般读者，最后是科学史家和科学哲学家。许多评论者都认为，第一类对象确可以从这部经过精心研究而写出的著作中受益匪浅。对第二类对象则显得过于艰深。第三类对象一定会深感兴趣，并引发热烈的争论。例如，法因在评论中说，曹和别人一样，并没有解决结构实在论所面临的困难。艾奇逊则对曹天予的自主存在的客体的本体论性质定义提出了质疑。确实，对实在论这种形而上学观点，是无法证实和证伪的。但是科学提供的大量科学理论成功的事例为科学实在论提供了“无奇迹”论证。曹的著作有力地表明，作出结构实在论的承诺，对现代物理学的研究，仍然是必

不可少并大有裨益的。

我的简略评介远远不能反映曹天予这部400多页著作的丰富内容。我认为曹的这部著作，对我国的理论物理学工作者和学生、科学哲学和科学史工作者、自然辩证法工作者，都是一部有关当代物理学史和物理学哲学、当代物理学家的自然观的很有价值、不可多得的参考书。我恳切地希望，这部著作能够早日在中国翻译出版。

范岱年

（原载：《自然辩证法研究》1998年第1期）

前　言

这部作品的目的是给出20世纪场论——从广义相对论到量子场论与规范场论——的一个广泛的综述。这些理论首先被看作是据以形成我们关于物理世界的概念的概念框架。本书的意图是给出关于场论的概念基础的一个历史批判的说明，从而考察这些概念在演化发展中的模式。 [xiii]

作为文化的重要组成部分，物理世界的概念涉及自然界的构成和运作方式的模型，包括关于物质的终极成分之间的基本相互作用的机制的假设以及关于空间和时间本质的解释。也就是说，这个概念涉及哲学家通常所谓的形而上学假设。现在，谈论形而上学已经不时髦了。这在科学研究的行业中尤其如此，目前大家首先关心的是局部的、基于经验的成功，以及社会利益和权力关系。在连观测事实的客观地位也受到社会建构论者挑战的时代，谁会在意弯曲时空或虚量子的本体论地位呢？但是，正如我们将在正文中看到的那样，形而上学的考虑对于在研究中另辟蹊径的物理学家是至关重要的。其

中的一个理由是这些考虑构成了他们的概念框架的基本组成部分。但是，目前形而上学在文化上的重要性比其对专业研究的贡献更深入也更广泛了。我自己的经历也可以说明这一点。

在阅读了笛卡儿(Descartes)、康德(Kant)、黑格尔(Hegel)、罗素(Russell)、爱因斯坦(Einstein)、海森伯(Heisenberg)和玻姆(David Bohm)的哲学著作以后，我开始研究理论物理学，我被物理学所吸引[xiv]纯粹出于文化好奇心，试图获得被物理学最新发展所认同的物理世界图景。我被告知牛顿图景是不合适的，机械世界观在19世纪已被实质上是场论世界图景的电磁世界观所取代。我也了解到，在20世纪，我们已经目睹了两次由相对论与量子论带来的物理科学中深刻的概念革命和本体论革命。物理学革命的结果是，探究物理世界基础的新概念框架把我们武装起来了。但是，由这些革命性理论所揭示的整合的世界图景是怎样的呢？在12年前，当我在英国剑桥大学开始从事科学史与科学哲学的研究时，我企图从20世纪的物理学家、物理学哲学家和物理学史学家的著作中寻找这样一幅图景，结果却劳而无功。

当然，我也从卡西勒(Ernst Cassirer)、石里克(Moritz Schlick)、赖辛巴赫(Hans Reichenbach)、波普尔(Karl Popper)、霍尔顿(Gerald Holton)、格林鲍姆(Adolf Grünbaum)、斯泰因(Howard Stein)、厄曼(John Earman)、施塔赫尔(John Stachel)、马丁·克莱因(Martin Klein)、库恩(Thomas Kuhn)、贝尔(John Bell)、希莫尼(Abner Shimony)、法因、雷德黑德，和其他许多学者那里学到了很多。例如，我已经知道，一些类似于普适性原理和对应原理的形而上学预设，在革命性理论的创立者的理论建构中起着重要的启发性作用。我也认识到，对于大多数受过教育的人们，这些理论的某些形

而上学意义，诸如机械以太的取消、平直时空的废除，以及在微观世界中个体事件的因果描述与时空描述的不可能性，已经作为我们的世界图景的重要部分被接受。但是，我并没有发现一幅整合的图景，更不用说对这幅图景的演化以及演化模式和方向的令人信服的说明。我决定填补这个空白，这部著作就是我努力的结果。

这部书首先是为理论物理学专业那些对他们学科中的基础问题感兴趣，并且力图从历史的视角来把握他们学科的内在逻辑和动力的学生写的。但是，我也尽力使这部作品让接受过基本科学教育的一般读者也能够理解，这些读者往往感到那些通俗作品无法满足他们对当代的自然概念的文化好奇心。我最后面对的读者是主流的科学史和科学哲学专家。虽然这部著作已经提供了这些学科的更深入的文化和社会学分析的基础，包含了许多哲学反思材料，但在目前的环境下，本书设定的方案不大可能让这些学者感兴趣甚至让他们接受。分歧来自不同的科学概念。反对目前立场的详细论证将在介绍性与总结性的章节中给出。这里，我只强调有争议的几点。 [xv]

对于很多从事科学研究的学者来说，任何根据由经验定律构成的科学理论所假设的终极成分和隐机制给出的世界图景的讨论，似乎等同于预设了建立在理论的不可观测实体和结构的基础之上的朴素实在论立场，而这简直是不可接受的。这种反实在论的立场具有悠久的传统。对于古典实证主义者来说，任何涉及诸如原子或场之类的不可观测事物的陈述，都超越了经验证据或逻辑推理的范围，因而是没有意义的，都必须从科学的论述中驱逐出去；因此，世界图景问题是个伪问题。对于生活在后经验主义时期的建构经验主义者或老练的工具主义者来说，描述假说性的不可观测事物的理论术语是允许的，但与存在状态不一致，因为这些术语只是拯救现象和作出预测的工

具，或是可观测事物的一种速记。这样，他们面临的问题是，这些工具的有效性的根源是什么。

为了回答这个问题，需要澄清工具与外部世界的关系。但是，工具主义者一直没有作出这样的澄清。不过，他们已通过借助于所谓的迪昂—奎因(Duhem-Quine)的非充分决定性(underdetermination)论题来怀疑理论术语的实在论解释，根据这个论题，没有任何一个理论术语能被经验证据唯一地确定。但是，这个论题的说服力完全是建立在把经验证据视为决定一个理论所假设的本体论的可接受性的单一判据的基础之上。一旦剥夺了经验证据的这种优先地位，那么把科学理论看作只是由经验的、数理逻辑的和约定的成分构成的简约主义观点就被更站得住脚的观点取代，其中形而上学成分(例如概念框架的可理解性与似真性)也被包含在其中并被视为理论可接受性的一个判据，这样一来就把科学理论放在了各时代累积的预设所构成的更广阔的网络和一种普遍性的文化氛围中，从而单单迪昂—奎因这一论题并不足以有力到让人怀疑理论术语的实在论解释。

更激进的是库恩的立场。如果迪昂—奎因论题全盘接受多种相互竞争的理论本体论——这些理论本体论全都与一组给定的证据相容，并平息了关于哪一个本体应当被视为真的争论——那么库恩(1970)则拒绝任何理论本体的实在性。他问道，正像科学史似乎已经向我们显示的那样，既然由科学理论假设的本体论总是被另一种不同的、经常是被后来的理论所假设的不相容的本体论所取代，而且不[xvi]存在本体论发展的连贯方向，那么我们如何能够把任何一种理论本体看作是世界的真实本体呢？而历史事实却是，一些假说性的本体论总是或明或暗地存在于理论科学中。因而库恩面临的问题是，为什么理论本体论在科学的理论结构中是如此地不可或缺。

库恩的工作已经产生了一些反响。不像那些孤傲地沉迷于科学理论的抽象逻辑和语义分析的逻辑经验主义者，库恩已经试图在对实际存在着的理论的历史考察的基础上发展他的科学观。但是，他的一些追随者对他把科学实践仅仅限制在概念方面的做法并不满意。他们大声疾呼实验和研究机构、社会利益和权力关系以及诸如此类东西的重要性。但是，据我看来，在这一点上库恩基本上是正确的：科学实践的核心在于理论建构和理论争论。实验是重要的，但是如果没有放入某个理论语境，它们的重要性就将是不可理解的。所有外在因素对于我们理解科学是有趣的和有价值的，但只有当它们与理论，即理论的产生、建构、接受、使用和结果存在关联时才是如此。否则，它们将与我们对科学的理解无关。在这一点上，福曼(Paul Forman)的著作(1971)是重要的，因为它描述了有助于接受在量子力学中发展起来的非因果性概念的德国文化氛围，虽然它没有触及在非因果性概念的形成中文化氛围是否扮演建构性角色的问题。

与此相似的问题被科学社会学强纲领的倡导者采纳，并作出了肯定的回答，他们持科学是一种社会建构的立场(见 D. Bloor，1976；B. Barnes，1977；A. Pickering，1984)。在通常的意义上，现在很少有人与他们争论科学的社会建构特点。但是在争论中真正使人感兴趣的一点是他们关于自然的特殊立场。如果自然被设想为在科学的建构中不起作用，那么社会建构论者将没有理论源泉来提出涉及科学理论的真理地位和客观性的问题，而相对主义和怀疑论将是不可避免的。但是，如果允许自然在科学的建构中起作用，那么科学将不单单是一种社会建构，而且社会建构论者在把科学的本质解释为关于自然的知识这方面将不会有什么成就。

科学研究的最近风尚追随着一种社会建构论者从后现代主义文化

[xvii] 的最新时尚中吸收了大量辞藻的激进版本。时尚追求者仅仅把科学看作是事实与知识的劝导、操纵和制造的修辞艺术；知识不过是一种权力冲动，与真理或客观性全然无关；而客观性只不过是一种意识形态，与科学知识实际上是如何制作的无关。他们争辩道，从事科学研究的学者的重要任务不是要找出谁发现了事实以及谁构造了概念与理论，而是谁控制了实验室；不是要在广义相对论所预言的红移能够真正被观测到的意义上来解释为什么科学是有效的，而是要询问谁从科学中获利。这些时尚追求者面临的问题是，他们只能彼此间讨论，而永远不能与科学家严肃地讨论他们的主要活动，即他们的理论活动。

另一种有影响的立场是普特南(Putnam)的内在实在论(1981)。这种立场允许我们谈论抽象实体、真理和实在，但只能在一个理论框架内进行。既然任何一种谈论总是在一种确定的框架内进行，人们似乎不可能回避这种立场。应当注意这种立场与卡尔纳普(Carnap)关于语义学框架的立场(1956)具有亲密关系。这两种立场都拒绝涉及独立于我们的语义学框架的理论实体的客观实在性这个外部问题。对这种立场的辩护，正如普特南所做的那样，存在于如下声明中：即使形而上学实在果真存在，我们也没有方法接近它。设想如果后继理论假设的本体论能够被证明彼此之间没有联系，那么这种立场与库恩的立场是不可分辨的。但是，如果在后继的理论中本体论承诺的演化具有一致的方向，那么意味着什么呢？因而，普特南不得不面对实在论者提出的老问题：理论演化方向一致的实体基础是什么？

因此，为了对关注自然终极成分和运作方式之根本假设的物理学概念史作出辩护，我们必须回答两个问题。首先，为什么这些形而上学假设对于物理学是必不可少的？其次，我们是否有接近形而上

学实在的方式？ 对第二个问题的肯定回答将在正文中给出。 这里是我对第一个问题所持立场的概述。

众所周知(见 E. A. Burtt, 1925；A. Koyré, 1965)，在中世纪末期，亚里士多德哲学衰落，具有毕达哥拉斯主义面孔的新柏拉图主义复兴。 后者把数学看作是实在的基础，而把宇宙看作其结构根本上是数学的。 人们假设可观测的现象必须符合数学结构，而数学结构应当对更进一步的观测和超越已知事实的反事实推理给出暗示。 从那时候起，就存在一种强烈的倾向，特别是在数理学家中间，即把数学结构看作是描述物理世界的基本实体及其行为的概念框架。 [xviii]

形而上学在同一个时期发生转变的另一个方面是，随着科学理性本身的勃兴，在因果性的概念中，目的因被动力因取代了，伴随着它的是权威的权力被理性的即因果性的推理的权力所取代。 于是，力而不是亚里士多德的目的因成为因果性的动因，并被看作是自然现象的形而上学基础。 在某种意义上，物理学的所有后继发展都能被看作是被寻求一种机械的或任何其他的模型所驱动的，即被寻求描述能够被理解为因果动因的力的模型所驱动。

这些变化的同时出现，导致了在 17 世纪由笛卡儿、玻意耳(Boyle)，在某种程度上还有牛顿(Newton)，为了解释和预测而在物理学中发展起来的假说-演绎法的兴起。 正是在物理理论的这种特定结构中，我们能够发现本体论假设的不可或缺性的深刻根源。 力、场、以太、刚性或动力学时空、虚量子、禁闭夸克、规范势，所有这些假说性的(在发展的特定阶段，它们被称为是形而上学的)实体对于理论物理学是不可或缺的，因为它们为历史上出现的、这个学科所固有的假说-演绎法所必需。 理论中一些终极本体的假设，提供了把实体的某一集合还原为另一个更简单的集合的基础，从而赋予这个理论

一种统一的效能。不充分注意理论物理学的理论结构的这种特征，就不可能对理论物理学及其效能作出合适的理解。在这方面，我认为迈耶森(Meyerson，1908)是对的，当时他坚持认为，现代科学作为一种自从哥白尼时代以来出现的建制，不过是自然形而上学的一个更深入的阶段，而常识则假设基于可观测现象的永恒实体的存在。

在本书中，场论的处理是高度选择性的。考虑到这门学科的丰富内容，只能如此做。这种选择一般来说被我的科学理论观所引导，尤其是被我对场论的理解所引导。这些材料提供了考察与解释各种论题的视角，因而在最大程度上确定了在这门学科的进展中的各种论题的意义。选择和解释材料的一般框架严重依赖于一些组织化的概念，诸如那些形而上学、本体论、物质、实在、因果性、解释、进步之类的概念。但是，这些概念在文献中经常是模糊不清的。为了澄清迷雾，我在第1章致力于详细解释我对这些概念的使用，并阐述一些有关方法论重要性的论题。故事的出发点，即从经典场论的兴起和危机到洛伦兹(Lorentz)的工作，将在第2章概述。正文的主体与我希望详细阐述的对发展的结构的理解相一致，被划分为三篇：几何纲领(geometrical programme)、量子场纲领(quantum field programme)和规范场纲领(gauge field programme)。每一篇由三章构成：前期历史，概念基础的形成，更进一步的发展与评价。理论发展的哲学含义，尤其是那些关于实在论和合理性的内容，将在总结性的章节中探讨。

[xix]

关于参考书目的说明：只有那些实际上用来为这本书做准备的著作才列在参考书目中。除了至关重要的原始著作之外，对原始著作提供解释的当代学术著作也列入了参考书目。但是，我没有试图对次级文献提供完整的参考书目；只有那些与我对这个主题的解释有直

接关联的著作才包括进来。关于在前面两章中频繁提到的现代知识的历史的一般背景，我简单地要求读者查阅一些著名的历史著作，而没有给出具体的原始文本的参考书目，而这事实上能够在所提到的著作中找到。

我分两阶段实施本书的写作计划。在第一个阶段(1983—1988年，英国剑桥大学期间)，我从与我的导师赫西(Mary Hesse)和雷德黑德，以及我在剑桥最亲密的朋友巴特菲尔德(Jeremy Butterfield)的许多讨论中受益良多。他们每一位都阅读了手稿的早期几个版本，而且作了大量评注和修改建议。我对他们的宝贵批评、帮助，以及最重要的鼓励深表感谢。我也感谢莫法特(Henry K. Moffatt)的关注、鼓励和帮助，以及伍德(David Wood)的友谊和帮助。

第二个阶段开始于1988年，是在我离开英国剑桥大学到了美国马萨诸塞州坎布里奇市以后。在过去的七年中，我非常幸运地有大量机会与施韦伯和科恩(Robert S. Cohen)讨论问题，我欠他们两位很多人情；我也与施塔赫尔和希莫尼有过一些详细的讨论。我对他们对当代物理学与哲学的知晓与理解印象深刻，并非常感激他们对全部或部分手稿的重要批评和建议。从20世纪80年代中叶以来，我已经从与布朗和库欣的长期友谊中获益匪浅，我对他们是非常感恩的。我非常感激哈曼(Peter Harman)的激励和鼓舞。我也感谢很多为明辨事 [xx] 理而与我亲切交谈的物理学家，他们当中有阿德勒(Stephen Adler)、巴丁(William Bardeen)、科尔曼(Sidney Coleman)、菲舍尔(Michael Fisher)、乔治(Howard Geogi)、格拉肖、格罗斯(David Gross)、贾基夫、约翰孙(Kenneth Johnson)、卡达诺夫(Leo Kadanoff)、洛(Francis Low)、南部阳一郎(Yoichiro Nambu)、波尔钦斯基(Joseph Polchinski)、特霍夫特(Gerardus 't Hooft)、韦尔特曼、温伯格

(Steven Weinberg)、怀特曼(Arthur Wightman)、威尔逊(Kenneth Wilson)、吴大峻和杨振宁。

本书写作的首批支持来自英国的海外研究学生(ORS)委员会的海外研究学生奖学金，剑桥大学的校长助学金，以及剑桥大学三一学院的海外研究学生费用助学金(1983—1985)的联合赞助；然后是三一学院的研究奖学金(1985—1990)，美国国家科学基金的赞助[赞助号DIR-No. 9014412(4-59070)](1990—1991)，以及布兰代斯大学的赞助(1991—1992)。没有这些慷慨的支持，我不可能完成这部著作。此外，在第二阶段，由于哈佛大学、布兰代斯大学、波士顿大学和麻省理工学院的加盟，我的工作效率已经大大提高了。对于所有这些研究机构，我都欠了很多人情。我尤其感谢哈佛大学的希伯特(Erwin N. Hiebert)和霍尔顿(Gerald Holton)，布兰代斯大学的施韦伯，波士顿大学的科恩，以及麻省理工学院迪布纳科技史研究所的布赫瓦尔德(Jed Buchwald)与西姆哈(Evelyn Simha)的热情。

我所欠人情最大的是我的家庭成员，他们在困难的条件下，对我追求学术上的卓越给予毫无保留的感情支持和实际支持，而且容忍我的“书呆子气”。对于这所有的一切以及其他许多事情，我感谢他们。

曹天予

波士顿大学

目录

第1章 导 论

在这部专著中，论题的处理是选择性的与解释性的，并且为一些哲学的和方法论的考虑所激发和指引，诸如那些集中围绕形而上学、因果性与本体论的概念，还有那些关于进步与研究纲领的概念。 但是，在文献中，这些概念往往是以模糊不清的方式表达出来，而且还导致了误解和争论。 关于这些动机的争论，涉及其实在论、相对主义、合理性与还原论的含义；在最近几年里，由于理论性论文的急剧的重新定向，争论甚至变得更激烈了。 因此，必须尽可能清楚地阐明在我所选择与解释的相关材料的框架中的这些东西。 我将在 1.1 节中通过叙述自己的一般科学观来开始这种努力。 在 1.2—1.4 节中详细阐述了涉及物理学的概念基础的论题之后，我将在 1.5 节中转向我对历史与科学史的理解。 导论将在 1.6 节中以主要情节的概述结束。 [1]

1.1 科学

近代科学作为一种社会建制，以一连串的人类实践出现在 16 与

17世纪。借助于科学，人们能够系统地理解、描述、解释与控制自然现象。在导致科学发生的重要因素中，我们发现有工艺（仪器、技能和行会或专业学会），社会需要（由新兴资本主义要求的技术革新），巫术和宗教。作为日常活动的一种延伸，实践层次上的科学旨在解决疑难、预测现象，以及控制环境。在这一点上，科学与工艺和社会需要的关联是毋庸置疑的。[1]

〔2〕 但是，作为一种满足人类对于所生活于其中的宇宙本质的好奇心的方式，一种满足他们希望拥有对物理世界的融会贯通的概念的方式（一种对世界的构造、结构、定律与运作方式的理解，不是根据表象来理解，而是根据实在，也就是根据世界的真实图景、终极原因与统一性来理解），与近代科学的诞生来说更相关的是巫术与宗教中的特定传统，即，正如耶茨（Frances Yates，1964）与默顿（Robert Merton，1938）分别指出的文艺复兴时期的赫耳墨斯主义（Renaissance Hermetism）与新教宗教改革（Protestant Reformation）。在这些传统中，理解、操纵与变革物理世界的可能性和方式被理性地予以讨论，并且通过诉诸物理世界的特定前概念（preconceptions）来加以辩护，这些前概念深深地扎根于人类思想中，但只有通过宗教改革与近代科学的兴起，才在现代思想中占主导地位。

在这些前概念中，最重要的是假定物理世界具有一种超验的品格。在炼金术传统中，人们设想一种宇宙和谐的异教宇宙论，其中神具有超自然的能力，而人类与这些超验实体分享着相似的性质与能力，并且能够和它们进行交流与互动。在宗教传统中，世界的超验性存在于上帝的意识之中，因为世界的这种存在是上帝意志的结果，而且自然的运作方式也是上帝设计的。这个超验假设，已经为近代科学同时具有神秘与理性的双重品格铺下基本含糊性的基础。超验

假设是神秘的，因为它旨在展现自然的秘密，自然的秘密要么与预先神秘地建立起来的宇宙和谐有关，要么与神圣的上帝有关。超验假设又是理性的，因为它设想自然的秘密能够被理性所接近，并为人类所理解。超验假设的理性主义含义，被新教神学家在他们的宇宙论中进行了详尽阐述。除了其他因素以外，这种阐述不仅对于近代科学的产生是关键的，而且它还给近代科学遗留了一些基本特征。

根据新教宇宙论，上帝通过自然来进行工作，根据他有意识设计的、规则的自然律来运作，因此自然律是确定的、不可改变的、必然的和彼此协调的。既然上帝的无限权力被设想为是通过有规则的途径来实现的，并反映在世界的日常事件中，因此人们相信井然有序的世界能够被科学家热情洋溢地进行研究，这些科学家借助于他们的经验的帮助，试图发现自然现象的原因与规则。新教神学的宇宙论原理(把上帝与自然现象及其定律联系起来)提供了研究自然界的宗教动机与合法理由。在加尔文主义的追随者看来，对自然界系统的、理性的与经验的研究是通向上帝的舟车，甚至是对赋予人类生命的上帝表达崇敬的最有效方式。其理由是，经过不断的探索研究与实施操作，自然将逐渐展现理性，不断逼近完美，并最终呈现上帝杰作的真实本质，彰显上帝的荣耀。这种超验的动机自近代理论科学出现以来，一直指引着其发展。而且这种超验动机的世俗化版本在当代科学文化中依然流行。[2] [3]

虽然近代科学对于理性地、系统地理解与改变世界的伦理的与感情的动机起源于与资本主义伦理(在经济、经营管理与政治学领域系统思考的产物)相一致的清教徒的价值观，但是近代科学，作为一种智力追求，主要是由修正了的古希腊原子论与重新发现的阿基米德(Archimedes)所塑造，尤其是由文艺复兴时期的新柏拉图主义所塑

造。后者与柏拉图的形而上学，或者与普遍和谐的神秘宇宙论及内在一致的体系联系在一起，其目的仍是借助于数学神秘主义与数学符号体系的帮助，对有关自然现象的人类经验作出理性的综合，并因此引起了哥白尼与开普勒(Kepler)的好奇心，点燃了他们的想象力，正如爱因斯坦、狄拉克(Dirac)、彭罗斯(Penrose)、霍金(Hawking)以及许多当代超弦物理学家的好奇心与想象力被激发起来一样。

另一个在文艺复兴时期的巫术中普遍流行的前概念与“自然的齐一性”(uniformity of nature)有关。持此观念的巫师相信，相同的原因总是导致相同的结果，而且人们只要施行与所设立的规则相一致的宗教仪式，想要的结果将不可避免地随之而来。虽然这种有关自然事件之间的联结关系的信念只有类比这样一种基础，但是这种信念显然是主张自然事件的进展是规则地确定的、并为不可改变的定律所决定的力学思想的先驱；在力学思想中，定律的运作能够被精确地预见和计算，而机遇与偶然的因素因此从自然的进程中被排除掉。

采用规则的定律划分自然的领域，有助于我们进一步把上帝从经验科学的因果性思想中驱除出去，把自然从超自然现象的领域中分离出来，并采用自然主义的原因作为解释自然现象的基础。与此相关的信念是自然力是可操作和可控制的。没有这种信念，就不会有用符号语言学操作的占星术、炼金术与神秘术的实践活动。数学符号[4]体系受到推崇，只是因为人们相信数学符号体系是操纵自然力、征服自然界的操作方式的关键。

人们可以非常有趣地注意到，在 16 与 17 世纪，科学处于形成期，神秘图景与科学图景共存和重叠，巫术与宗教的前概念有助于塑造科学的特征，诸如：(i)除了客观性，还有理性主义与经验主义，都与新教宇宙论设想的自然的超验品格有关；(ii)基于“自然的齐一

性”思想的因果推理；(iii)科学的理论表述中的数学符号体系；以及(iv)在科学的实验精神中显示的操作意愿。

但是，和巫术与宗教不同，科学拥有颇具特色的工具来承担其任务：理解与操作世界。在这些工具中重要的是：(i)专业学会与出版物；(ii)理性的批评以及建立在怀疑精神与求同存异基础上的争论；(iii)经验观察与实验、逻辑、数学(数学的系统应用通向特定的证明模式)；以及最重要的(iv)富有成果的隐喻、概念框架与模型，并利用这种理论结构，世界的结构与运作方式能够被接近、描述与理解。

一个科学理论必定具有一些在有可证伪结果的意义上的经验性陈述，也必定具有一些虽然不可单独被证伪、但对理解与解释现象却是重要的假说性陈述。假说性陈述是用理论术语表达的：既有抽象原理，又有不可观测的实体与机制。在科学哲学家中间，对于理论术语在理论中作为组织经验的启发性手段的功能，没有什么争议。有争议的是理论术语的本体论地位：我们应该以实在的方式处理它们吗？对于感觉材料经验主义(sense data empiricism)，答案绝对是否定的。但是，孔德(Comte)反对波动说的声明，马赫(Mach)反对原子论的立场，被证明是严重的错误。对于约定主义(conventionalism)，科学的基础比感觉材料更广阔，在逻辑与数学之外还包括约定。尤其是，它把约定的数学表达看作是理性的基础，可观测量与不可观测量都从属于它。但是“这些建构性的约定其真实性程度如何？”是约定主义者不愿也不能回答的问题。对于内在实在论(internal realism)，理论术语的实在性能够被接受，但是只有在这些术语出现的理论中才能被接受，而不能从理论中分离。理由是，即使形而上学的实在性存在，我们也无法接近。我对内在实在

论的不满，在本书中随处可见。

[5] 在最后一章涉及理论术语的本体论地位时，我将捍卫的是结构实在论(structural realism)的立场。简单地说，这种立场坚持认为，在成功理论中的结构关系(经常直接用数学结构来表达，但也能用模型与类比来间接表达)应当被视为是真的，不可观测实体的实在性被逐渐建构，而且在一种理想的情形下，最后被这些结构关系以唯一的方式决定。

对结构实在论立场的直接反对意见是，认为这是一种令人讨厌的现象主义(phenomenalism)，其中可观测量的经验真理被其数学真理所取代，而且没有给不可观测量的实在性留有余地。批评者将会争辩道，给出一个解释远比仅仅写出一组方程来概括观测到的规则更为困难。

为了预知我在最后一章的论证，只需指出，除了可观测量的结构关系以外，也存在对于理解与解释更重要的不可观测量的结构关系。针对反对意见(任何这样的结构关系必定受到不可观测实体给予的本体论支持)，我的回答是，在任何一种解释中，虽然结构关系在它们是可检验的这一意义上是实在的，但涉及结构关系的不可观测实体的概念总是具有一些约定的成分，而实体的实在性是在它们所涉及的越来越多的关系中建构或推导出来的。一旦我们接受了结构对实体的本体论优先地位，我们在适应不断变化的实体解释时，就会具有更多的灵活性。

1.2 形而上学

形而上学，正如我理解的那样，是由关于宇宙的终极结构的预设构成的。首先，形而上学涉及关于世界实际上是由什么构成的，或

者说世界的基本本体实际上是什么的问题。 世界是由客体、性质、关系或过程构成的吗？ 如果我们把客体看作基本本体，那么问题就接踵而来：客体的范畴有哪些呢？ 存在既是精神的也是物质的客体吗？物理客体的基本形式是什么——粒子、场或其他形式？ 此外，空间与时间的本质是什么？ 本体论讨论的一个核心难题涉及实在的划界标准，因为形而上学总是与实在的或基本的或原初的实体联系在一起，而不是与副现象或派生物联系在一起。 诸如笛卡儿与莱布尼茨(Leibniz)那样的近代哲学家，对这个问题的经典回答是，只有物质实体是实在的。 一个物质实体不借助于任何其他物质实体就能永恒 [6] 地自存，它在没有任何外因的情况下就有活动的能力。 但是，正如我们将看到的那样，可以存在基于潜能、结构或过程而不是物质实体的其他实在概念。

其次，形而上学也涉及控制世界的基本实体的原理。 例如，存在着同一性原理；这个原理宣称，个体应当能够在时间中变化并同时保持自身的同一。 类似地，连续性原理宣称没有一种变化是间断的。 有很多其他形而上学原理在科学理论的建构中扮演着重要的调节性或启发性的角色，诸如简单性原理、统一性原理，以及时空的可图像化原理。 但是，这些原理中最重要的是因果性原理，它被设想为支配自然的运作，并有助于使实体的行为成为可理解的。

因此，通过诉诸本体论的假设与调节性的原理，形而上学提供了科学争论的前提与似真性支持，并且非常不同于有关观测现象及其规则的经验的、实用的与局部的陈述。 形而上学在传统上是高度猜测性的。 也就是说，形而上学的断言是尚未检验的预设，也不必是经验上可检验的。 但是，这些具有认识论与本体论意义的预设在文化中被如此牢固地确立起来，以至于它们对科学家来说，显得就像常识

一样的直觉。既然这些确立起来的假设给出了世界的似真图景，而且事实上决定了认同这些假设的人们的思维方式的深层结构，因此形而上学构成了文化的一个重要部分。

休厄尔(William Whewell)曾经指出：

> 物理发现者不同于贫乏的思辨家，不是因为在他们的头脑中没有形而上学，而是因为他们有好的形而上学，而他们对手的形而上学是坏的；应该把物理发现者的形而上学与他们的物理学结合在一起，而不是把两者割裂。
>
> (Whewell, 1847)

正如我们在正文中将会看到的那样，形而上学假设能够用物理参量来使自己骨肉丰满。但是，比这更重要的是，形而上学提供了一种在其中某些理论能够被提出并进行检验的可理解的概念框架。众所周知，古代科学最初是从形而上学的思辨中发展起来的。但是，甚至直到现在，科学仍然与由形而上学思想提供的这样或那样的世界图景相联系。一种现象的解释总是根据特定的世界图景给出的。哪一种本体，物质、场、能量或时空能最好地解释现象，这个问题对于物理理论极具重要，远比经验定律的细节重要。[7] 例如，牛顿力学的经验内容只是细微地被爱因斯坦相对论所修正，但是没有人否认这是物理学发展中的一大步，因为关于欧几里得空间、绝对时间与绝对同时性、超距作用的老思想被扫到了一边，并且世界图景因此发生了变化。

物理学与形而上学之间的相互作用的例子是很多的。在从 19 世纪的电磁理论到狭义相对论与广义相对论的发展中，考虑到相对性原理的普适性，形而上学假设的引导功能被广泛承认。另一方面，物

理学的发展，特别是量子力学、量子场论与规范场论的发展，也具有深刻的形而上学内涵，并且强烈地改变了我们关于物理世界的概念，这点我们将在本书主体部分中加以讨论。物理学与形而上学之间相互渗透的必然结果就是这样。不仅形而上学对于物理学研究是不可或缺的，而且物理学也给我们提供了一条直接通向形而上学实在性的道路。例如，阿哈罗诺夫—玻姆效应(Aharonov-Bohm effect)* 与贝尔不等式(Bell inequality)的实验研究已经分别大大地澄清了量子势的本体论地位与量子态的本质，这两者都曾经被设想为是不可接近的形而上学问题。出于这个理由，希莫尼(1978)把这种研究称为是实验形而上学(experimental metaphysics)，强调它在检验形而上学假设的过程中重要的物理作用。

因此，仅仅把形而上学作为科学理论中排除了经验内容与逻辑结构的一个可接受的残余物是不恰当的。相反，在形而上学提供了物理实在的基本模型从而使得理论成为可理解的意义上，而且也在形而上学在因果性的确定概念的基础上偏爱特定的解释类型的意义上，形而上学具有特殊的科学内容。在动力因的力学概念被视为解释的基础的情况下，形而上学假设不仅决定了物理理论的假说-演绎结构，而且也使得固有的还原方法论成为必要。更进一步，既然实证主义者在涉及原因的时候成为不可知论者，而且只有实在论者认真地看待原因，那么实在论的形而上学假设的固有含义也不应当被忽视。

1.3 因果性

随着近代科学的兴起，权威或传统被解释现象的原因所取代。科 [8]

---

* 简称 AB 效应。——译者

学的终极目的之一就是理解世界，而这是通过科学解释，也就是通过发现各类现象的原因来接近的。但是，根据亚里士多德(Aristotle)，存在着不同类型的原因：质料因、形式因、动力因和目的因。在近代科学兴起以前，建立在目的因基础上的目的论解释，是解释的主导模式。随着新柏拉图主义、阿基米德主义以及原子论在文艺复兴时期的兴起，开始了科学解释的基本假设的转换。例如，哥白尼、开普勒、伽利略(Galileo)与笛卡儿相信世界的基本真理与普遍和谐能够用简单精确的数学表达式完美地表示。自然的数学化在一定程度上导致了形式因的普及。但是，在反对目的论解释的过程中，最普遍、最有力的因果性概念，却是建立在动力因基础上的力学概念。与目的因和形式因不同的是，动力因的思想集中在原因如何达到结果的问题上，也就是这种传导的模式问题上。根据机械论，因果性能够被约化为时间和空间中物体的运动规律，而且可观测的定性变化能够为不可观测的组分微粒纯粹的定量变化所解释。

力学解释具有不同的变种。根据笛卡儿，宇宙是广延的充盈物，不存在真空，任何给定的物体是连续地与其他物体相接触的，因而宇宙的若干组成部分的运动只能通过直接的碰撞与挤压来互相传递，超距作用是不可能的。不必举出伽利略曾用力或吸引来说明特定种类的运动，更不用说开普勒的“活力”了。所有发生的事情都符合一台无摩擦运转的机器的规则性、精确性与必然性。但是，根据牛顿，力是运动的因和果，尽管力自身不得不通过运动定律来定义。对于牛顿，以及对于惠更斯(Huygens)与莱布尼茨来说，因果性的可理解性主要定位在力的概念中。于是，力的传输的具体机制立即成为严重问题。这个问题对于物理学的后继发展如此重要，以至于它实际上规定了物理学发展的内在逻辑。寻找这个问题的答案的

努力导致了场论、量子场论，最后是规范场论的诞生。

存在着不同形式的力学解释。第一种形式是，自然现象可以根据在现象中实际涉及的物质粒子的排列来解释，根据在这些粒子中间起作用的力来解释。在第二种形式中，采用一些力学模型来表示现象。这些模型未必被视为是表征实在的，只是被视为证明现象在原则上能够用力学来描述。也就是说，这些力学构造使得现象成为可理解的。第三种形式是，力学解释也可以表述为拉格朗日分析动力学的抽象形式体系。由此得到的运动方程独立于力学系统的细节，但是现象依然以质量、能量与运动等力学术语来解释，因而仍可归到这个形式体系所包含的力学解释的原理之中，尽管它们并没有用特定的可图像化的力学模型表示出来。 [9]

在这三种形式中，模型的使用具有特殊的重要性。甚至分析动力学的抽象形式体系，也需要通过模型来阐释。而且，既然物理研究的主要动机之一在于，当现象层次上的直接原因不能解释时，就去寻找在基础层次上力的因素，那么包含假说性的、不可观测实体与机制的模型的假定就是不可避免的。因此，假说的必要性内在地包含在力学解释的思想本身中，或者在对动力因的追求中。

任何一个假说必须与自然的基本定律相容，与所有正在讨论的现象有关的、得到普遍接受的假设一致。但是，一个假说只有与基本定律及普遍假设相联系，才能证明它是正确的。因此，假说的特定内容必须加以调整，以允许关于所研究的现象的陈述的演绎。但是，关于不可观测量的假说是如何能够解释现象的呢？假说是如何能够加以调整，以便这个目的能够实现？对这些问题，本书最后一章在结构实在论的基础上的尝试性回答是，只有当从日常经验或其他已知现象中抽象出来的类比模型(任何假说都是一种模型)的结构相似

于现象的结构时，一个假说才能完成它的解释功能。

物理理论的假说-演绎结构具有直接的形而上学含义：如果具有一系列不可观测实体的一组相互一致的假说能够充当现象世界的原因，那么似乎不可否认的是，假说世界给出了实在世界的真实图景，而且现象世界能够还原为这个实在世界。例如，大多数力学解释都设想一个具有不可观测的运动原子或基本粒子这样一些隐蔽本体的真实世界，作为物理实在底下的基石。也存在着其他可能性。例如，莱布尼茨把力的密集连续统作为现象的形而上学基础。18与19世[10]纪的其他物理学家超越了力学解释，但仍然在假说-演绎体系的一般框架内工作，提出了不同的非力学本体论，诸如活力原理、火、能量与力场等。[3]在每一种不同的本体论中，物理学家不仅提供了不同的物理理论或研究纲领，而且也提供了现象世界背后的实在世界的不同概念。

1.4 本体论

与现象或副现象相反，也与纯粹启发性和约定性的策略相反，本体论(ontology)作为在实在的逻辑构造中不可归纳的概念要素，是同实际存在有关的，即同与任何外在事物无关的自主存在有关。既然一种本体论给出了关于世界的一幅图景，那么它就充当着理论赖以建立的基础。这有助于解释在科学的理论结构中本体论的演绎与构造性的作用。

虽然本体这一术语经常指称实体(substance)，例如就机械论世界观而言，其中基本本体是运动着的粒子，但情况未必都是如此。本体的概念，甚至在终极的真实实在的意义上，也比实体的概念更广泛，实体的概念又依次比实际存在物(entity)与个体(individual)的概

念更广泛。例如，人们能够争辩说，如同开普勒那样的新柏拉图主义者愿意做的那样，数学关系式代表着宇宙的结构，是实在的基础；甚至作为因果原理的力，也不得不根据数学关系式来定义。然而人们能够争辩道，从考察物理实在的结构视角出发，任何一种数学结构必须被实际存在物之间的物理关系所支撑，如果数学结构不仅仅是一种空洞的名称，那就只能以涉及它的关系来定义。这只是卡西勒所谓的“表征实在的功能模态”的一个例子。另一个例子能够在怀特海(Whitehead)的过程哲学中发现。根据怀特海，活动功能不是一种不变的基底质料的功能；确切地说，物理对象是一种联系，一种关于基本功能的或多或少有点永恒的模式。他论证道，自然是进化过程的一种结构，实在是一种过程，而物质的东西从比事物更基本的活动和生成过程中流溢出来。

当然，这是一个非常有争议的问题。迈耶(Julius Mayer)追随莱布尼茨，把力看作是自然的基本作用，力作为自然活动的具体化身，应当被看作是非机械的物质实体。而对迈耶森(Meyerson)来说，实际存在物对于解释是重要的，不应当消融在关系或过程中。更重要 [11] 的是，作为历史事实，本体的概念几乎总是与实体的概念相联系。这种联系构成了物理科学教程的基础，而且不可在物理学基础的检验中忽略掉。

那么，什么是实体呢？实体总是由本质的或基本的特性的群体来刻画的。这些特性存在于空间与时间中，在空间与时间区域的变化中守恒，而且所有其他特性都能够还原为这些特性。既然实在的本质只能根据符号表示来讨论，那么本体在一般的意义上，实体在特殊的意义上，作为实在的一种模型，是科学本身的一个片段，不能从科学中分离出来。因此，对什么是基本特性，在不同的理论中有不

同的理解，而且每一种理论决定了自身的一种实体。但是，自从莱布尼茨时代以来，普遍持有的假设认为实体必定是基本的（与副现象相反）、活动的或是活动的源泉，是自存的，即实体的存在不依赖于任何其他事物的存在。本书的主要结论之一是，一般来说，概念革命把原来的实体转变为一种副现象，因此改变了我们关于世界的基本本体是什么的概念。

在经典物理学中，笛卡儿把空间或广延看作实体。牛顿的情况复杂得多了。除了实体，他的本体论中还包括力与空间。并且他的实体不仅指被动的物质粒子，而且也指主动的以太。对于莱布尼茨，实体是基本活动的中心。这种活动不是材料或物质（matter）的显现，而是活动本身就是实体，物质只是这种活动的外表。

在莱布尼茨以后，占主导地位的观点是把实体看作是内在活动的对象，通常划分出不同的本体论范畴：离散的个体（诸如可见的有质量粒子与不可见的原子）与连续的充盈物（诸如笛卡儿的广延物与经典场）。个体是空间上受约束的对象，而且至少还有一些其他的性质。它通常被刻画为能够被识别、被重新识别，并可以与这个领域的其他成员相区分的东西。[4]这里同一性（identity）是由基本特性的守恒来保证的，而可区分性（distinguishability）在预设了对象的空间约束的不可入性（impenetrability）上有其根源。个体的概念通常是与粒子的概念相联系的，因为两者都必须是离散的，但是由于可区分性与不可入性的要求，个体的概念要比粒子的概念狭窄些。在量子理论中，量子粒子可看作是全同粒子，但它们既不能被重新识别，也不能从其同类粒子中区分出来。因此，它们不是个体，但仍然能被解释为粒子，主要是因为静止质量、电荷与自旋的守恒。

[12]

这是理论依赖于我们关于实体的概念的一个例子。另一个有趣

的例子是能量的本体论地位的概念。在传统上，能量被设想为是实体的最重要特征，因为它显示其载体是活动的，而作为作用能力的量度，它是守恒的。但是，作为一种可度量的性质而不是自存的客体，能量本身通常不被视为是实体。例如，当诺伊曼(Carl Neumann)声称势能是基本的、而且能够自身传播时，麦克斯韦则坚持能量只能存在于与物质实体的联系当中。[5]出于同样的理由，根据唯能论，能量作为纯粹的能动性是物理实在的基础，这通常被指责为现象主义，因为它拒斥实体。但是，唯能论也能以另外的方式解释。如果能量被看作具有始终活跃着的新特征的实体之类的东西，当它的量保持守恒的时候，它总是在改变着自己的形式，那么这种东西究竟是什么呢？看来唯能论似乎将成为詹姆斯(James)的功能主义与怀特海的过程哲学的先驱。

这两个例子暗示着一个本体论假设不仅对于特定的理论是基本的，而且对于一个研究纲领也是基本的。让我们从这样一个视角出发，对场论纲领的产生作更细致的考察。电磁场被视为通过空间连续传导的电磁力的原因。在19世纪的物理学中，场的实体性是一个有争议的论题。有时人们论证道，麦克斯韦确立了场的实体性，因为他证明了能量存在于场当中。但是，这种论断是有问题的。对于麦克斯韦，场不是一种客体，而只是遵循牛顿运动定律的机械以太的一种状态。这意味着对于麦克斯韦来说，场不是自存的，因而不可能是物质的。能量在场里面的存在所确立的只是以太的实体性，而非场的实体性。

有时候人们也会论证道，机械以太的去除使得场的实体性的去除成为必要，因而这就支持了时空点是场论的基本本体的论断。[6]但是，根据我的意见，正是机械以太的去除，才建立了场的非物质的实

体性。这方面的理由是，在这种情形下，场变成了场能唯一可能的储存器，而场能预设了一种实体作为它的储存器。至于时空点，为什么它们不能被视为场论的基本本体的理由是，在前广义相对论的框架中，不可能论证时空点是活动的或者是活动之源；而在广义相对论中，时空点不是自存的，因为它们总是被引力场占据，而且更重要的是被引力场个体化。[7]根据我的意见，连续的实体场是世界的基本本体的假设必须被视为场论的第一个基本信条，虽然在物理学史上并不总是这样的。

[13]

场通过其连续性与个体区分开来，这与个体的离散性形成对照，并且场通过不同部分之间的重叠性与个体之间的不可入性形成对照。诚然，场通过边界条件引入的周期性也能显示为离散的一种形式（见6.5节）。但是，这种离散存在不同于个体的离散存在，个体的离散存在是永恒的，而场的离散存在则是短暂的。[8]

当电磁场被看作是世界的基本本体，而不只是一种数学装置或机械以太的一种状态时，存在着比这表面差别更深刻的东西。[9]场是一种新型的实体，莱布尼茨的原始力可视为它的先驱。场因其非机械的行为而既不同于物质个体，也不同于机械以太。这种新的非机械本体的引入，开创了一个新的纲领——场纲领。场纲领不同于机械纲领之处在于它的新本体与通过场传播作用的新模态。洛伦兹电子理论的出现，在两种意义上标志着机械纲领的终结。首先，它抛弃了机械以太，致使电磁相互作用的传输不能在机械纲领中得到解释。其次，它引入了一种独立的、不能还原为机械本体的电磁场实体。而这就为场纲领的进一步发展铺平了道路（见第2章）。

人们并不否认作为科学研究的概念基础的本体论假设的重要性。但是，库恩主义者会论证道，被一个理论所假设的任何本体总是被一

个不同的、为后来的理论所假设的、而且经常是相反的本体所取代；这一历史事实似乎令人信服地显示了我们的本体论假设与真实世界无关。但是，正如我将在最后一章中论证的那样，本体论的历史性更替，至少在20世纪场论的语境中，不是没有模式与方向的。即一种旧的本体总是被证明是一种副现象，能够从一种新的、更基本的本体中导出。而这显然给理论本体的实在论解释提供了支持。

有两条限制应当直接附加上去。首先，这种模式并不总是以直线发展的方式实现，而经常是通过辩证的综合来实现。其次，本体的还原只是科学发展的一个维度。正如我将在11.4节规范场论的语境中表明的那样，因为客观涌现的存在，世界的不同层次具有它们相对自主的本体，这个自主本体不能还原为终极基质。这要求本体论的多元化。[14]

1.5 历史与科学史

历史不仅是过去事件的集合，而且是由具有原因、模式与方向的运动所构成。如果我们考虑埋藏于过去事件中的大量信息，历史编纂就不能只是对真实历史的复印，而必须是有选择的。材料是根据其历史意义而被选择的，而历史意义是由历史学家的解释决定的。既然我们对过去的解释是根据现在和我们正在挺进的未来而逐渐定型的，而且也因为历史编纂是从流传下来的传统的东西着手，这意味着总结过去的教训而进入未来，历史编纂是过去与现在之间的对话，其目的是根据现在去反思过去，并根据过去来理解现在与未来，即试图理解从传统到未来的转变。

因此，一篇好的史评必须提供一种合乎情理的或说服力强的有关历史运动的原因、模式与方向的假说，以使过去的事件成为可理解

的，而一种全面的或建设性的展望能够扩大我们对历史运动的理解，因而打开通向进一步探究的道路。既然一些原因只是偶然的，另一些却是普遍的，那么对一个我们用以理解过去事件的假说来说，核心的问题是“什么应当被视为主要的原因？”在科学史的情形中，这是有争议的一点。社会建构论者与理性的科学思想史学家在这方面的分歧已使他们的关系有点紧张。

的确，科学作为文化的一种形式，不可能从社会中分离出来。首先，科学使用的语言是社会交往的产物。其次，科学试图解决的问题，解决问题的动机、材料与技术资源都是由社会提供的。第三，为科学研究运筹资源的机构，大体上是受社会经济结构支持与约束的。所有这些考虑都指向科学活动的社会特征。而且，科学家只能从适应当前情形的文化环境中得到解决问题和解释结果的思想、隐喻与概念框架。这种环境给予科学知识一种明确的社会、文化与历史的特殊性。在这种肤浅的层次上，对科学的社会特征几乎没有异议。[15]

在科学的社会建构论说明中，“社会的”指的是科学家之间的社会关系，大体上也指科学家与社会之间的关系，包括科学家团体、科研机构与社会经济结构。“社会的”也指文化的非科学形式，诸如宗教与巫术。但是，社会建构论与众不同之处在于，它是在与个体和知识分子的相对照上来定义社会的。但是，在社会建构论的非平庸的意义上，它对科学活动的说明是不完备的、存有偏见和具有严重缺陷的。建构论说明存在缺陷的原因是，忽略甚至故意否认所有科学活动受追求自然知识的目的严格约束与持续指引这个事实。实际上，自然知识的目的是建立在把科学作为特定的社会建制的概念基础之上的。忽视了这个关键点，科学活动的社会建构论说明就不能以

它现在的方式被接受。出于这个理由，我把概括了人类持续的智力追求的概念历史，而不是社会历史，作为科学史的主体，虽然我也承认在概念历史的基础上，社会历史可能是有趣的，如果它成功地提供了科学作为一种智力追求实际上如何发展的更完备的图景的话。

有时候，物理学史主要是根据它经验的成功方面来编纂的。但是，它也可以用另外的方式来编纂。基础物理学的发展必然涉及本体论假设的巨大变化。物理学史显示了大多数伟大的物理学家从事研究的目的是为了寻求真实的世界图景，这个图景为概念革命铺平了道路。这不是物理学成长的非主要特征，而是其最重大进展的核心特征。因此，物理学史在某种意义上是世界观表达的历史，其核心由镶嵌在描述性框架中的本体论假设所构成。更具体地说，我打算在这本著作中要做的是，通过对 20 世纪场论的历史作批判说明，来探寻由这些理论所提示的变化的世界观的模式、方向及其演化；即揭示在这些理论中已经发生的本体论转变与本体论综合。

1.6 主要情节的概要

在我的说明中，除了法拉第(Faraday)把场看作空间的状态的推测可以作为智力上的先驱之外，场纲领在把场看作世界的基本本体的意义上，是从洛伦兹的电子论开始的，而在爱因斯坦的狭义相对论出现以后才被广泛接受。但是，这两种理论都预设了空间点或时空点的独立本体作为场的支撑。所以，在还原论的意义上，这个阶段的场论不是完备的纲领。

[16]

下一步我将在场纲领的发展中展开主要的情节，即在爱因斯坦的广义相对论(GTR)中展开情节。在广义相对论中，传递相互作用的引力场，与时空的几何结构不可分离地关联在一起。出于这个理

由，我把广义相对论以及在这个方向的后继发展称为几何纲领。

广义相对论的解释是有争议的论题。它依赖于是否把物质、场或时空点看作它的基本本体；它也依赖于对物质、场与时空之间关系的理解。在历史上，存在三种解释。首先，爱因斯坦本人追随马赫，把可称重物体(ponderable bodies)看作完全决定引力场与时空的几何结构的唯一的物理实在。引力场方程的真空解的发现使得这个解释站不住脚(见4.3节)。尔后，外尔(Hermann Weyl)与爱丁顿(Arthur Eddington)发展了一种观点。根据这种观点，时空的几何结构被视为物理实在，而引力场可约化为这个几何结构。按照这种观点，引力被解释为时空流形曲率的一种显现，即引力是几何化的。我把这种观点称为强几何纲领(见5.2节)。爱因斯坦自己从不欣赏这个纲领。最后，爱因斯坦在统一场论中把引力场看作代表终极物理实在的整体场的一部分，时空是整体场的结构属性。在某种意义上，这是爱因斯坦的时空几何引力化的延伸，这种立场起源于爱因斯坦在1915年为广义协变性原理作辩护而提出的点重合论证(point-coincidence argument)。我把这称为弱几何纲领(见4.2节，5.1—5.3节)。虽然数学形式体系保持不变，但在不同的解释中，本体的优先性是不同的。

广义相对论构成了几何纲领的出发点。在这个纲领的进一步发展中，人们作出了四个推广：(i)引力与电磁学的统一；(ii)通过引入时空流形的扭转，质量与自旋效应的统一；(iii)引力与量子效应的统一；(iv)物质与场的统一。这些推广大多数是在强几何纲领下作出的。但是，我将证明，在强几何纲领中所作出的努力，在把量子效应整合起来的时候，必然会导致这个纲领的崩溃(见5.3节)。这就证明了我的主张：强几何纲领是不合适的。但是，弱几何纲领在尝

[17]

试把它的范围延伸到电磁相互作用、弱相互作用与强相互作用时，也是不成功的。

再者，几何纲领具有严重缺陷，其种子就埋在广义相对论自身之中；即广义相对论的方程，在膨胀宇宙的大爆炸或坍缩的恒星或黑洞中，都不可能避免奇点或奇异解。奇点的不可避免暗示着广义相对论在充分强的引力场的情况下必定失效。但是，存在着几何纲领的另一种推广，它与量子场论框架中规范理论的新近发展有联系，而且通过这种方式，能克服奇点困难。这个有希望的尝试将在第三篇讨论。

20 世纪物理学中场纲领的另一个重要变体，始于约当(Jordan)与狄拉克的量子电动力学，即量子场纲领。具有悖论色彩的是，场纲领似乎被量子理论挖了墙脚。首先，量子理论对整个空间的能量连续分布设置了限制，而这与场本体论相矛盾。其次，它也违背了可分离性原理，根据这个原理，具有零相互作用能量的远距离系统应当是物理上互相独立的。最后，在量子理论中，对于粒子的量子跃迁或创生与湮灭过程，不给粒子具有连续的时空路径留有余地，而这是与在场纲领中传导相互作用的模式相矛盾的。所以，我怎样才能声称量子场论(QFT)应当被看作场纲领的一个变体呢？对这个问题的回答依赖于我对基本本体论和量子场论中传导作用的模式的解释。

那么，什么是量子理论的本体呢？总的来说这个问题很难回答，因为非相对论性量子力学中的情形不同于量子场论中的情形。在非相对论性量子力学情形中，这个问题与波函数的解释紧密相关。德布罗意(de Broglie)与薛定谔(Schrödinger)坚持波函数的实在论解释。他们假设场本体论，拒绝粒子本体论，因为，他们论证道，遵守量子统计的量子粒子显示出不可鉴别性，因而不是经典可观测的个

体。但是，这种实在论观点遇到了严重困难：在多体问题中，它隐含着非物理的多维性、非实在的态叠加(诸如著名的薛定谔猫)，以及到现在还不能作任何解释的测量过程中的波函数坍缩。这些困难显示着经典场对于非相对论性量子力学的本体来说，不是合适的候选者。[18]

玻恩(Max Born)在概率解释中拒绝波函数的实在性，剥夺了波函数的能量与动量，而且假设了一种粒子本体论。但是，玻恩并不主张经典的粒子本体论。因为存在着一些困难，诸如量子统计与双缝干涉实验，阻碍了他这样做。作为对这些困难的回应，海森伯把波函数解释为一种势(见 7.1 节)。人们应当注意，这种对实在论的场本体论的有意义的妥协*，在粒子本体论的概率解释面临两难困境时扎下了根，而且为量子场论中本体论的巨大转变铺平了道路。

总之，非相对论性量子力学的情况就是这样。除了一部分不能作实在论解释的数学工具以外，基于场本体论的数学结构(波动方程)与基于粒子本体论的物理解释(概率解释)的概念之间不相容的混合，使得想用经典粒子或经典场的术语为这一理论寻找一种连贯的本体论显得特别困难。有两条出路可能是有希望的。对待这一困难的工具主义解决方案是，把所有理论实际存在物当作只是为了经济地描述观测到的现象的相似构造，而不是量子世界真实面目的忠实描述。但这只是一种逃避，并不能解决量子理论中的解释问题。或者，人们可以试图重塑自己的实体概念，不把本体局限于经典粒子与经典场的二分法。在 7.3 节中将考察量子场论早期历史中重塑实体概念的过程，并可概述如下。

---

* 指势等于新的物理实在的观点。——译者

随着费米子场量子化的引入，对于实体的重新分类来说意义重大的本体论转变就出现了。在那以前，对于多体问题有两类量子化程序：二次量子化(second quantization)与场量子化(field quantization)。这些程序各自预设了粒子本体与场本体。二次量子化程序实际上与场本体无关，它只是粒子的量子系统中的一种表象变换。在这种表述中，粒子是永恒的，它们的创生与湮灭不过是不同态之间的一种量子跃迁现象，而概率场不过是用于计算的数学工具。与此相反，场量子化程序是从由场谐振子集合体表示的实在场出发，通过采用产生与湮灭算符来显示场的粒子性质，而这些算符能够被解释为场量子的激发与退激发。

[19] 但是，在非相对论性量子力学中，当把量子化条件应用到场谐振子的时候，被量子化的东西只是场的运动(能量、动量等)，而不是场本身。把场转变为粒子的一种集合体，实际上是没有问题的，这将使得粒子本体论假设成为必要，除非人们设想运动的量子化使得运动载体（即场本身）的量子化成为必要。但是，这是一个意义重大的形而上学假设，要求对我们的实体概念作出激进的改变。

这两种量子化程序并不是矛盾的，而是平行的，它们从不同的本体论假设开始。人们饶有兴趣地注意到，这两种程序的始作俑者之一狄拉克在他的早期著作(1927b，c)中，已经把两者相互合并，虽然他显示出一种粒子本体论的倾向，而把场还原为粒子的集合体。这种倾向暗示着上面所提到的激进形而上学假设实质上也是由狄拉克作出的，虽然只是无意识的。

由约当与维格纳(Wigner)引入的费米子场量子化，在量子场论的解释方面开始了更激进的改变。首先，一种实在论的解释取代了波函数的概率解释：他们表述的波函数必须被解释为一种实体场，否

则，作为场量子的粒子不可能从场中获得它们的实体性，而薛定谔的实在论解释面临的一些问题仍然没有解决。其次，场本体论取代了粒子本体论：物质粒子(费米子)不再被看作一种永恒的独立存在，而是作为场的一种短暂激发，即场量子。因而这验证了我的主张：量子场论开启了场纲领的一个主要变体，即量子场纲领。

但是，实在场观点的复归，显然在实在模型的逻辑中留下了一条鸿沟，摧毁了实体的传统概念。人们引入了一种本体论的新概念。这种新的本体不能还原为经典粒子本体的概念，因为场量子缺少一种永恒的存在与个体性。它也不能还原为经典场本体的概念，因为量子化的场已经失去了它的连续存在。20世纪的场论似乎暗示着量子化的场，这些场与一些非线性场(诸如孤波)一起构成了一种新型的本体论，雷德黑德称之为瞬息(ephemeral)本体论。在量子场论的奠基者当中，约当是自觉地推进这种新的本体论，并重塑了实体的概念，而狄拉克与海森伯则是不自觉地这样做了。

量子场论的新本体论具体体现在狄拉克真空概念中。作为一种[20]植根于量子激发与重正化的概念框架的本体论背景，狄拉克真空对于诸如魏斯科普夫(Weisskopf)的电子自能计算与丹科夫(Dancoff)对散射的相对论性修正的讨论之类的计算，是重要的。存在于狄拉克真空中的涨落强烈显示出真空必定是某种实体性东西，而不是空虚的。另一方面，根据狭义相对论，真空必定是零能量与零动量的洛伦兹不变态。考虑到在现代物理学中能量不太严格地被看成对实体来说是根本性的，这似乎意味着真空不可能被看成是一种实体。这里我们陷入了一种深刻的本体论的两难困境，暗示着改变我们关于实体之概念和作为实体性质的能量之概念的必要性。

在量子场论中，传导相互作用的方式在两个方面不同于经典场纲

领。首先，相互作用通过场量子之间的局域耦合来实现，这里耦合的精确含义是量子的创生与湮灭。其次，作用不是通过连续的场来传递的，而是通过局域地耦合于实粒子、并在它们之间传播的离散虚粒子来传递的。[10]因此，量子场论中相互作用的描述通过局域耦合的概念深深地植根于算符场的定域激发(localized excitation)概念中。

但是，由于不确定关系，局域激发 (local excitation) 必定使得任意数量的动量都是可利用的。因此，定域激发的结果将不仅仅是一个单一的动量量子，而必须是动量量子的所有适当组合的叠加。这就具有意味深远的结果。首先，相互作用不是由单个虚动量量子(费恩曼图中用内线来表示)来传递，而是通过虚量子的无穷多数目的适当组合的叠加来传递的。这是量子场论中场本体论基本假设的必然结果。其次，具有任意高动量的无穷多虚量子，导致它们与实量子相互作用得出的无穷大贡献。这就是著名的发散困难(divergence difficulty)。因此，只要这个严重的困难没有解决，量子场论就不能被视为一个自洽的理论。在历史上，这个困难最早是通过重正化程序绕过的。

原初的重正化程序的本质是把无穷大量吸收到质量与电荷的理论参量中去。这等价于把以局域激发概念为基础的精确点模型模糊化。虽然量子电动力学满足可重正化性的要求，但关于弱相互作用的费米理论与关于强核力的介子理论却不满足这个要求。不过，这个困难能够通过使用规范不变性(gauge invariance)的思想来克服。规范不变性是确定基本相互作用形式的一般原理，它以在量子场纲领 [21] 中针对基本相互作用而发展起来的新纲领——规范场纲领——为基础。

规范不变性要求引入规范势，具有规范势的量子承担着传递相互

作用和弥补在不同时空点内部自由度的附加变化的责任。在规范理论中规范势扮演的角色与广义相对论中引力势扮演的角色是相当的。广义相对论中的引力势是与一种几何结构相联系的（切丛中的线性联络），而规范势是与相似类型的几何结构，即主纤维丛的联络相联系的。广义相对论与规范理论的理论结构之间深刻的相似性，暗示着规范理论本质上也可能是几何的。

基本物理学的新近发展[超引力与现代卡卢察—克莱因理论(Kaluza-Klein theory)]已经打开了联系规范势与时空额外维上的几何结构的大门（见 11.2 节与 11.3 节）。因此，如果我们把规范场纲领表达为这样的形式：相互作用是通过与要么存在于内部空间、要么存在于时空的额外维的与一种几何结构不可分离地相关联的量子化规范场实现的（规范场量子与物质场相耦合，并在相互作用的传递中起作用），那么把规范场纲领看作是几何纲领与量子场纲领的综合看来是合理的。

人们饶有兴趣地注意到，只有在规范场纲领把两者进行综合之后，量子场纲领的本体与几何纲领的本体的密切联系才成为可辨明的。这个事实暗示着，综合的概念有助于人们认识到理论本体在经历概念革命之后其结构性质的连续性。科学思想的综合需要将以往科学思想进行转换。当我们把综合的概念延伸到理论体系的基本本体的讨论时，我们发现本体的综合也需要以往本体概念的转换，而且作为一般特征，使实体变为一种副现象，并伴随着基本本体的转变（见 12.4 节）。

这个特征暗示着本体论综合的概念已经抓住概念革命的一些特点，概念革命的结果是在新本体论基础上新研究纲领的诞生。按照这种观点，把革命前纲领的旧本体直接整合到革命后纲领的新本体中

去是非常不可靠的。不过，一些已经发现的世界的结构关系，诸如体现在旧本体论中的外部对称性与内部对称性、几何化、量子化等等，显然将会在经历科学革命之后仍然保持着。我揣测，尽管研究领域是可扩大的，然而这里仍然有着朝向真实的局域结构的本体发展的一贯趋势。因此，概念革命的发生，绝不意味着理论科学的发展如同库恩设想的那样是根本不连续的或不可通约的。相反，在一些特定种类的科学实在论中，概念革命是连续的、可通约的与相容的。 [22]

另一方面，通过从20世纪场论的历史分析中抽取出本体论综合的新概念，我们也发现科学的成长未必采取连续积累的单线发展形式，因此避免了站不住脚的“收敛实在论”(convergent realism)。本体论综合的思想作为世界结构的连续性与积累的辩证形式，较之在解释概念革命的机制与科学进步的模式中的收敛思想更加强有力。既然科学合理性的思想存在于得到世界的实在结构的更多知识的意图中，那么在科学成长的综合观点看来，概念革命只是实现这种合理性的一种方式。

由这种综合观所提出的基础研究的未来方向，不同于由库恩的不可通约性观点所提出的发展方向。对于一些库恩主义者来说，科学研究的方向主要是由社会因素所支配的，很少涉及智力因素。对于其他学派，一组智力因素在特定的范式中是重要的，但在新范式中要么不起任何作用，要么甚至是不可通约的。但是，按照综合观，则必须尽可能多地保留先前理论的内在结构。

由综合观作出的对未来研究的建议，也不同于单线发展观所作出的建议。根据单线发展观，现存的成功理论必须取作未来发展的一个模型。但是，按照综合观，既然超越现存的概念框架的新综合总是可能的，那么应当劝告科学家要对各种可能性保持开放的心态。

从这个角度来看场论的未来发展，很可能它的未来并不是单独依赖于规范场纲领内部的研究，即试图把希格斯机制与汤川耦合整合到场纲领中，而且也依赖于对S矩阵理论(它的基本思想——诸如本体论与力的本质的基本思想——根本不同于场纲领的基本思想)产生的结果的使用。

[23] **注释**

1. 见 Zilsel(1942)。
2. 例如，见 Gross(1992)和 Weinberg(1992)。
3. 见 Cantor and Hodge(1981)。
4. 见 Strawson(1950)。
5. 能量的本体论地位的上述解释将在第7章的讨论中起重要作用。
6. 例如，见 Redhead(1983)。
7. 关于在本体论上引力场优于时空点，一些论证将在4.2节中给出。
8. 在一些非线性场论中，诸如孤波理论，这种差别被模糊了，从而不可入性变成了对个体性最重要的判据。
9. 这首先出现于洛伦兹电子理论中，其中以太没有机械性质，并且与绝对静止空间同义，而在爱因斯坦的狭义相对论中，以太的概念只是简单地从其理论结构中消除。
10. 虚粒子的传播只是一种隐喻，因为一个可识别虚粒子的确定的空间轨迹从未被探测到。虚粒子的本体论地位是一个难题。一些预备性的讨论能够在布朗与哈里 (Brown and Harré) 的论文(1988)中发现。把虚粒子的传播概念化这个困难，已经被费恩曼通过其路径积分表述所掩盖(见 Feynman and Hibbs，1965)。

第 2 章
经典场论的兴起

虽然我打算探究的发展是从爱因斯坦的广义相对论开始的，然而，如果不经过合适的历史审视，那么既很难把握广义相对论发展的内在动力，也很难把握作为场纲领更深阶段的后继发展。对场纲领本身的兴起的适当说明，能够恰如其分地提供这样一种审视。这一章的目的就是提供一种说明，其中简要概括了导致场纲领兴起的主要动机与发展的根本假设。[1] [24]

2.1 力学框架中的物理作用

如同我们在第 1 章中提到的，在近代初期出现的两种智力倾向，即关于世界的机械化(mechanization)与数学化(mathematization)，有效地改变了人们关于实在与因果性的概念。根据笛卡儿与玻意耳等机械论哲学家的观点，物理世界不过是运动着的物质。根据开普勒与莫尔(Henry More)等新柏拉图主义者的观点，物理世界在结构上是数学的。作为两者的综合，物理世界的内部实在，表现为只是

其运动由数学定律支配的物质体。这里，物质既可以采取充盈的形式，就像对笛卡儿来说的那样；也可以采取微粒的形态，就像对伽桑狄(Gassendi)、玻意耳与牛顿来说的那样。正如我们马上将看到的那样，这两个力学体系之间的差别导致了对物理作用的不同理解。

世界的机械化也暗示着，人们能够在物质体的空间运动中发现现象的真实本性、所有变化与效应的本质与原因。例如，玻意耳论证道，颜色的本质不过是分子的位移。但是，开普勒受新柏拉图主义[25] 鼓舞，坚持认为：现象的原因——“它们为何如其所是”的理由——存在于基本的数学结构中；凡是数学上为真的，也必定是现实中为真的。因果性概念中的差异，即动力因与形式因的差别，也导致了对物理作用的不同理解。正如我们将看到的那样，这一点与源于对不同力学体系的不同理解纠缠在一起。

我们感兴趣的中心论题是如何解释诸如引力、电磁之类的物理作用，它们表观上是以超距作用传递的。在笛卡儿体系中，既然不存在虚空，表观的超距作用不得不以某些精致的物质形式、某些以太形式的实体(诸如火、气)以及各种流溢物通过碰撞或挤压作为传递介质。这种力学解释的典型例子是笛卡儿的磁理论。根据笛卡儿，精致物质的流溢物穿越磁体与周围空间在一个封闭的圈中循环。既然流溢物的流动导致磁体与铁之间的空气稀薄化，它们就在外界空气的压力下一起受力。按照这种方式，笛卡儿就把表观上超距传递的磁作用，还原为物质粒子运动导致的接触作用(contact action)。

在原子论体系中，原子是被虚空隔开的；如同玻意耳所论证的那样，仍然有可能借助于相互作用物体接受与施加的碰撞或挤压所引起的介质实体的局部运动，用接触作用解释表观超距作用。但是，从逻辑上讲，承认虚空使得接受超距作用作为一种实在的可能性成为必

要。在历史上，超距作用的概念被广泛接受的主要原因是因为牛顿引力理论享有天文学上的成功经验，而建立在接触作用概念基础上的笛卡儿的引力涡旋理论却失败了，这一点不久就被惠更斯认识到。根据基本的接触作用来解释诸如内聚力、弹性与磁性之类物理现象的困难，反过来也有助于人们接受超距作用。

但是，另一个重要因素也对超距作用的广泛接受作出了贡献，那就是新柏拉图主义者培育起来的因果性的形式概念。根据这个概念，如果某些特定现象能够被归入到特定的力的数学定律下面，那么它们就得到了解释，因而也是可理解的与真实的。因此，牛顿第二运动定律的方程 $F = ma$ 似乎建立了因果关系，即力是物体产生加速度的原因。既然数学定律被看作是现象的真正原因，那就没有必要探究力的起因。牛顿的一些门徒与追随者，特别是英国经验主义哲学家洛克(Locke)、贝克莱(Berkeley)与休谟(Hume)，非常严肃地看待这个概念。牛顿自己想得很深，感到物体具有作用于它们所不在的地方的能力是不可思议的，虽然有时他也倾向于把这种因果性的形式概念作为避难所。 [26]

超距作用的概念受到莱布尼茨的批评。通过诉诸连续性的形而上学原理(据此，原因与结果是连续地互相联系的)，莱布尼茨拒绝了微粒在虚空中运动的思想，因为这意味着物体密度在其边界不连续变化。对于莱布尼茨来说，物体从不会自然地运动，除非它被另一个物体碰撞或挤压，而引力的神秘本质只有通过精致物体的推动，才会以一种可说明的方式发生。

莱布尼茨对超距作用的批评被博斯科维克(Roger Boscovich)所拒绝；具有讽刺意味的是，他自己对莱布尼茨的反对意见却也是建立在连续性原理基础上的。博斯科维克论证道，如果物质的终极粒子

是有限的，那么在它们的边界，密度将有一个不连续的变化；而且如果它们相互接触，它们的速度将会不连续地变化，从而需要一个无穷大的力。无需求助于弹性部分的无穷回归，博斯科维克就得出结论，物质的本原必然是无广延的简单点，具有相互施加力的能力，力的大小则由它们之间的距离所决定。也就是说，物体之间的碰撞最后必定涉及能够用粒子之间距离的连续函数表示的超距力。

如何解释博斯科维克的力函数是一个有趣的问题。它只是一种抽象的关系吗？或者我们应该把它看作表征连续的实在吗？如果只是一种关系，那么它是如何转化为一种实际存在物，以至于能够在空间中独立传播呢？如果它表征着一种连续的实在，那么它的本质与构成是什么，我们如何能够描绘穿过这种连续介质的作用的传递呢？这些问题是有趣的，不仅因为博斯科维克的思想(由具有力学性质的数学点来取代广延的粒子，并把力看作某种独立于物质粒子而存在的准实体)的内在优点，也因为这些思想影响了后来的哲学家与物理学家，比如康德与法拉第。至于博斯科维克本人，他恰好稳稳地处在数理的实证主义传统中。他只是根据描述运动的变化的数学函数来定义力，而没有任何检验力的终极本质的意向。对于博斯科维克，正如对于牛顿，力既不代表任何特殊的作用模式，也不代表任何神秘的特性，而只是有质量物体的一种接近与后退的倾向。

[27] 2.2 连续介质

虽然植根于新柏拉图主义的形式因概念，被经验主义与实证主义的新哲学倾向加强了，但是对表观超距作用的动力因的探究仍持续进行下去。这种研究在机械世界观方面有其根源。从终极意义上讲，力学解释需要识别构成机械论基础的永恒实体；在机械论中，原因被

一步步有效地传递给结果。这种不可抗拒的强烈愿望，能够在牛顿对以太的沉思中找到范例。

以太思想具有悠久的历史，牛顿并不是始作俑者。以太思想对于机械世界观是如此重要，以至于从机械论哲学的早期兴起开始，它就在文献中变得非常流行。这方面的理由是，以太实体能够为穿越距离传递运动提供介质，而且更重要的是，一类具有非机械性质的以太流体能够用于解释各种非机械现象，诸如内聚力、热、光、电、磁与引力。例如，笛卡儿与他的追随者建议用几种流溢物来解释光、电和磁现象。

作为传递作用或非力学效应的传递介质的精致介质，各种以太实体(诸如光和引力以太，电流体、磁流体与热质流体，火与燃素，它们的分布或状态的变化会在普通物体中导致可观测的变化)被设想为具有不同于机械论的性质：它们是纯净的、不可见的、无形的，能够渗入到稠密物体中并弥漫于整个空间。通过求助于这样一个以太概念，牛顿试图给出引力的因果解释。没有求助于笛卡儿的通过邻近粒子施加的压力与冲力的概念，牛顿就把引力解释成为通过稀薄介质中的粒子施加的排斥力效应，这些稀薄介质中的粒子不均匀地散布于引力物体之内，穿越虚空，通过排斥力作用于可称重物体。

牛顿把以太看作是引力的原因与吸引的起因，从对自然的类比中推导出他的以太思想：类似于光线，牛顿认为以太流体也是由硬的与不可入的粒子形成。但是鉴于物体能够没有阻力地在以太中运动与漂浮，以太粒子，作为物质精致形式的组分，被设想为更稀薄、更精致、更有弹性，没有惰性，按照定律施加作用，以区别于接触作用的粒子。然而，这些特征不能赋予牛顿的以太非机械的资格，因为牛顿的以太仍然是由物质粒子构成的。此外，引力的超距作用只是被 [28]

同样在把以太粒子隔开的一定距离上起作用的排斥力所取代。

不过，在牛顿的以太中有一个重要的特征，标志着对机械论哲学的背离。那就是，尽管牛顿的以太由本质上具有惯性的物质粒子构成，但牛顿的以太具有活性的排斥力，不能还原为惯性。因此，如同康德认识到的那样，作为引力的一种活性起因，或者更一般地说，作为一种活性实体，牛顿的以太是力的构成的形而上学本原。

隐埋在牛顿理论著作中的作为活性实体的以太思想，被莱布尼茨更清楚地表达出来了。对于莱布尼茨，整个世界，普通物体以及光线，都包含着稀薄流体，即精致物质。甚至空间也不乏精致物质、流溢物或非物质实体。所有这些从终极的意义上讲，是由力的充盈(force plenum)构成的。力的充盈是一种活性实体，因为力作为运动的原因把活性具体化了。因此，对于莱布尼茨来说，自然的活力天然地包孕于力的充盈中，而所有这些细节都是一个动力学统一体的方方面面。如同海曼(Peter Heimann)与麦圭尔(J.E. McGuire)所提示的那样[2]，在从笛卡儿到玻意耳、牛顿、莱布尼茨与后来的科学家的发展中，我们发现了在实体本体论中惰性范围扩张的衰落与活力的兴起。这就是为什么我把活性(activity)看作是近代科学中实体定义的一部分的原因。

但是，一个我们已经在牛顿的以太中遇到的深刻困难是，任何一种连续介质，不管是否具有活性，如果它是由物质粒子构成，都不能提供替代超距作用模式的传递模式的本体论基础。正如康德指出的，其理由是，任何具有明晰边界的粒子，总是作用于自身以外的某种东西，因而就作用于它所不在的地方。因此，超距作用的真正替代者需要一种连续介质的新概念。

除了关于力的充盈、火或流溢物的一些猜测之外，对于这样一种

连续介质的概念，直到1802年托马斯·杨(Thomas Young)提出光的波动说，才成为科学上重要的主张。在波动说中，人们为一种跨越距离传递作用的新模式设想了一种弹性流体作为介质：作用通过流体的波动一步步地传递。不久，这种光以太在1816年菲涅耳(Fresnel)的表述中，就采取弹性固体的形式。

首先，光以太仍然被设想为由力或粒子的中心点，以及在它们之间的作用力所组成(Navier，1821；Cauchy，1828)，因而作用的连续 [29] 传递被还原为邻近的中心或粒子之间的相互作用。因为康德的论证在这里可用，于是超距作用不是被一种真实的替代品所取代，而只不过是转化到较低的层次而已。

既不是由无维度的点构成，也不是由广延的却具有明晰边界的刚体构成，这一事实上连续的与可变形的介质概念，是由斯托克斯(George Stokes)与威廉·汤姆孙(William Thomson)* 于19世纪40年代在光以太的数学研究中发展起来的。在他们新的连续介质概念中，我们能够在其中设想介质是可分的，介质的无穷小部分之间的作用，是由通过整个介质而非部分介质张力的重新分布来传递；而力的概念被作用于介质的无穷小部分的压强与应力的概念取代了。刻画连续介质数学特性的，是用来计算连续介质无穷小部分的变化的那些纯形式程序，而不涉及介质体系的隐秘关联。对他们的研究起关键作用的概念是势的概念，而势的概念早就由拉普拉斯(Laplace)、泊松(Poisson)与格林(Green)发展起来，与在连续介质中传递作用密切相关。

起初，拉普拉斯引力理论中的势，或者泊松静电学中某种物质的

* 即开尔文勋爵。——译者

势，不是定义为空间的一种实际的物理性质，而只是一种潜在性质：如果试验质量(或试验电荷)被引入到空间中的一点，那么就会有力作用于空间中的那一点。即引力势(或静电势)除了对被引入到定义了势的空间中的质量(或电荷)施加吸引力的潜在性质之外，不具有任何物理性质。存在着处理势的两种方式。如果用积分形式处理势，其中要对粒子之间所有超距作用做积分，一个势不显示通过物理上连续的介质来连续传递作用，而只是作为一种特设性工具，由此空间中每一点的力能够被定义为源的超距作用的合力。这样，势的概念似乎与连续介质的研究无关。

但是，势也可以是一种满足某个偏微分方程的量。既然首先由牛顿发明的微分定理把系统的邻近态直接联系起来，它们进入物理学就很适合用来表达施加于连续介质的邻近部分之间的作用，在偏微分方程中，就代表着连续系统的演化。更明确地说，在微分方程 $\nabla^2 V = 4\pi\rho$ 中，任何一点的邻域中一个势 V 的二阶导数只是以确定的方式与该点的密度有关，而在该点的势的值与距离该点一定距离的任何点的 ρ 值之间却没有关联。因此人们推想，势的概念能够转化为具有物理实在性的连续介质的概念。这种连续介质将为传递作用——既为超距作用也为接触作用——的替代模式提供一个物理基础。在这个模式中，介质的每一个点都由代表存在于该处的物理量(例如能量)的某个数学函数来刻画，而表观超距作用被解释为介质中应力状态的变化。

[30]

把连续势转变为这样一种物理上连续的介质，是与寻求表观超距作用的动力因的目的相协调的。但是，如果物理介质事实上是不连续的，如何能够把被设想为连续的数学工具作为这种物理介质的表征呢？一种可能是将不连续性看作在物理过程中被消除了。虽然这在

物理学上是可想象的，但在形而上学层面上不令人满意。它也没有指出，存在着一种数学模型，它能够处理严格连续介质中的作用，并且满足连续性原理。

19 世纪 40 年代的一些数理学家以不同的方式处理了这个问题。例如，斯托克斯不把物质的连续性或流动性或弹性看作组分粒子的宏观效应，而看作只在完全连续实体中达到极致的条件。他也从完全连续实体的方程出发，把普通连续介质的运动看作是可决定的。这里，我们发现新柏拉图主义实在概念中的结合点："凡是数学上为真的，也必定是现实中为真的"。这个结合点发挥了作用，并帮助数理学家把他们的连续实体的新颖概念定型。在概念上，这种迁移对于场论的发展具有深刻的后果。

2.3 机械以太场论

19 世纪场论的兴起是各种平行发展的结果。其中最重要的首先是光的波动说的成功，这主要归功于托马斯·杨与菲涅耳的工作，这个理论假设了一种弹性光以太作为解释光学现象的本体论基础；其次，由法拉第开创的，被威廉·汤姆孙、麦克斯韦与其他许多物理学家详细阐明的电与磁的研究，引导人们认识到光以太也可能是电磁现象的活动场所，因而在电磁理论中发挥了解释的作用，这种作用类似于它在光的波动说中所起的作用。

[31] 电磁现象能够非常成功地在超距作用的框架中得到描述。在法拉第与麦克斯韦的工作之前，较早的例子可以在法国物理学家库仑(Coulomb)、泊松与安培(Ampère)的系统表述中找到；较后的例子则可以在德国物理学家韦伯(Weber)与亥姆霍兹(Helmholtz)的系统表述中找到。因此发展不同框架的动机，主要是出于解释而非描述方

面的；即目的是发现表观超距作用的根本原因，并根据连续介质中作用的传递来解释超距作用。

但是，这样一种尝试在物理学史上已存在很久，显然远在法拉第与麦克斯韦登上历史舞台之前就出现了。如果我们把场的概念看作是所考虑空间区域的表示，而关于在其中各处运动的试验物体的势行为，场就被当作传递作用的动因，或者被当作一种其局部压力与应力是试验物体行为的原因的介质，那么正如海尔布伦(John Heilbron, 1981)指出的，18 世纪 80 年代的电学家已经有了"场"的概念，尽管没有用"场"这个术语。

然而，在 18 世纪与 19 世纪早期发展起来的场的概念，本质上是特设性工具，而不是物理实在的表示。理由是明显的：场只能设想为没有其他性质的物体作用的场所，因而不是独立的存在。考虑到这一点，我们发现源于法拉第的发展的新颖之处在于，把特设性工具转变为物理实在的表示。法拉第通过论证与实验演示表明，除了具有与物体的势行为有关的性质之外，介质还具有其他性质，其中最重要的是连续性与能量。

电磁场的概念作为一种新型的连续介质，首先是从流体，特别是与弹性介质的类比中导出的。1844 年，法拉第修正了博斯科维克关于物质粒子是力的收敛线的点中心的假说，而且提出在邻近粒子之间的电磁作用不是超距传递的，而是通过力线传递的。他进一步主张，电磁现象能够用这些力线来表征，这些力线的方向代表着力的方向。例如，与泊松势的思想相对照，法拉第对博斯科维克思想的修正，显然是电磁场论的第一个概念：尽管泊松势完全由粒子与电荷的排列决定，因此能够从电磁理论的描写中消除，但法拉第主张连续力线是一种独立存在，通过这些力线，电磁力能够连续地传递。[32]

首先，法拉第作出论证，反对把这些抽象的力线等同于一些充满空间的弹性介质，诸如由精致粒子构成的光以太。然而，受到认真看待光以太的威廉·汤姆孙的影响，法拉第对以太思想采取更同情的态度。他愿意把磁力看作是以太的功能，虽然他承认自己对力线存在与传递作用的实际物理过程没有了解清楚。有一段时间，法拉第对于力线的本体论地位不是很有把握：它们可能是某一点力的强度和方向的心智表征，或是一种物理实在，要么作为“纯粹空间的状态”，要么作为“我们称之为以太”的不可称重实体介质的状态(Faraday，1844)。

为了使连续力线成为一种物理实在，法拉第提出了四条判据。也就是必须满足：(i)在其发生作用的空间中，物质的存在会修正力线；(ii)力线独立于它们所终止的物体；(iii)力线在时间中传播；(iv)力线显示作用的有限容量，这相当于假设能量被储存在发生作用的空间中。最后一条判据是力线的连续性与实在性的必要条件，或者说是场(力线后来被称为场)的连续性与实在性的必要条件。人们不难看到，如果假设了运动的连续性，那么场必定拥有一些能量：如果把粒子放在发生作用的空间中原先没有物质的一点上，那么粒子可以获得一定的动能，而这些能量必须来自周围的空间；否则，将违背能量守恒(如果作用是以有限时间在物体之间传递，并且它的能量在这段时间间隔中不出现在发生作用的空间中，那么能量在这段时间间隔中就是不守恒的)。所有这些判据将引导后来的场概念的演变，即由威廉·汤姆孙，特别是麦克斯韦实现的从一种特设性工具到物理实在的表示的演变。

没有威廉·汤姆孙的先驱工作，麦克斯韦将不能前进。通过数学类比，威廉·汤姆孙首先在1845年建立了电磁现象与傅里叶

(Fourier)热传导工作之间的形式等价性，而后在1847年他又建立了静电学、电磁学与弹性固体的平衡条件之间的形式等价性。他用弹性固体的一般平衡方程的纯应变情形下的斯托克斯解来代表静电作用，而用纯旋转情形下的解代表电磁作用。在1849年和1850年，[33] 威廉·汤姆孙更系统地应用了这条推理线索，并且论证了磁学理论能够通过研究他称之为场的磁性物质的连续分布来加以发展。

首先，威廉·汤姆孙没有具体规定场的细节，只是暗示，基于法拉第的磁致旋光发现，磁力线可以等同于光以太，一种旋转的有弹性的连续介质。从这种机械以太的思想出发，威廉·汤姆孙提出了一系列理论，其中以太是一种基于自然界统一性的动力学实体，并且自身分化为自然界的各种粒子。从19世纪50年代开始，威廉·汤姆孙开始对他的数学结果表示的隐机制进行猜想感兴趣。1856年，他给出了1847年的形式结果的力学表示。在得知1858年亥姆霍兹关于涡旋运动的工作(这项工作表明，一定的涡旋运动具有某种永恒性)时，威廉·汤姆孙论证道，假设以太是穿过空间的完美流体，这些永恒的流体涡旋环可以等同于普通的原子。这是第一个把离散的原子与连续的充盈联系起来，用以太解释物质，避免粒子的超距作用或接触作用的模型。

威廉·汤姆孙有关以太的思想在19世纪下半叶特别有影响。然而，他提出的所有模型都具有机械的特征。尽管他关于连续介质与势的工作极大地帮助了麦克斯韦，但是威廉·汤姆孙自己并没有正确地理解麦克斯韦的电磁场理论。1884年，威廉·汤姆孙声称：

> 在能够作出一个事物的机械模型之前，我从不感到满意。如果我能作出一个机械模型，我就能理解它。只要我不能以任何方

式作出机械模型，我就不能理解：这就是我不理解电磁理论的原因。

(Thomson, 1884)

麦克斯韦从威廉·汤姆孙的数学类比出发，但他超越了数学类比，并且进入物理类比。从电磁现象与连续现象的相似性出发，麦克斯韦论证了电磁力一定包含通过连续介质的作用。为了发展他的电磁场理论，麦克斯韦把连续介质力学与势理论中的形式等价性问题的解，翻译为电磁学的语言。从作为结果的波动方程组，麦克斯韦推出了“光是由引起电磁现象的同一介质的横波构成”的结论(1861/1862)。然而，电磁介质与光介质的等同，仍然不能算是光的电磁场理论，因为构成光波的“横波”并没有借助电磁变量给出任何明确的解释。麦克斯韦完成的不是把光还原为电与磁，而是把两者同时还原为单一以太的作用机制。 [34]

麦克斯韦的以太思想像威廉·汤姆孙的一样，在性质上也是完全机械的。在他1861/1862年《论物理力线》(*On physical lines of force*)的论文中，麦克斯韦成功地把法拉第的思想具体化。法拉第设想在一个机械模型里面，借助一串物质粒子的极限情况，来表达力通过距离的连续传递。不久以后，也就是从1864年开始，麦克斯韦重新处理了对机械以太的这种具体说明。他1864年的论文《电磁场的动力学理论》(*A dynamical theory of the electromagnetic field*)，是他从设计具体模型转向采用抽象的广义的(拉格朗日)动力学方法的转折点。为了避免任何涉及以太的具体机制的不被允许的假设，并且仍然坚持他对机械世界观的承诺，麦克斯韦利用拉格朗日变分原理推导出著名的波动方程(1873)。他感到这个动力学理论是机械的，因为

牛顿力学可通过使用拉格朗日形式体系而得到重新表述。

在某种意义上，麦克斯韦是正确的。他对所发现的定律的解释是机械的，因为他的以太是牛顿动力学体系。如果基本动力系统被视为非牛顿的，那么就可能存在着麦克斯韦定律的非机械解释。但是，麦克斯韦对场概念的这种激进的修正还没有准备。虽然他给出了电磁作用的场论概念的数学表述，但是他没有完全明确地表达场的概念，因为他的场只是被设想为一种未知的动力学系统——以太——的状态。当他具体描述时，他的以太表现为极化粒子的电介质的极限。场作为这样一种机械实体的状态，既不是在不可还原意义上的存在，也不是在与物质相等意义上的存在，因此场没有独立的本体论地位。

人们经常声称，麦克斯韦对场论的最重要贡献之一是，他以波动方程建立起了场能的实在性。他真的曾经说过“不过，在谈论场能时，我希望被准确地理解。”其方程组的解显示了在传递电磁作用时出现时间滞后，这隐含着一些物理过程一定是在相互干涉的空间中发生的，这也是真实的。因此，法拉第的力线代表的是在这个相互干涉的空间中实际存在的某种东西。但是，这是否意味着麦克斯韦把场看作是承载把能量从一个物体传递到另一个物体的本体论上独立的物理实在？远远不是。关于所传播的东西，麦克斯韦说道：

[35] 我们不能想象在时间中的传播，除非要么作为物质实体通过空间的漂移，要么作为运动状态或已存在于空间中的介质中的应力的传播。

(1873)

关于能量储存于何处，他说道：

> 我们如何能够想象这种能量存在于空间的一点上，既不与这个粒子也不与那个粒子重合？实际上，每当能量在时间中从一个物体传递到另一个物体时，必定存在一种介质或实体，能量在离开一个物体之后又没有到达另一个物体之前，存在于其中。
>
> (1873)

这个论证的目的是要显示，发生传播与储存能量的介质是必不可少的。因此，麦克斯韦著名的“能量论证”建立的与其说是场能的实在性，还不如说是以太能量的实在性。

毋庸置疑，麦克斯韦关于作为场的载体的机械介质的概念是与对场的另一种理解不相容的。对场的另一种理解是，把“场设想为以非常独立于机械介质的方式，从一个粒子向另一个粒子发射”(Maxwell，1873)。这是诺伊曼于1868年在势能理论中提出的。麦克斯韦的论证矛头直接指向诺伊曼的场观点。诺伊曼的场概念超越了机械以太场的框架，并很好地处在力—能场概念的传统中。

在德国物理学中，诺伊曼不是第一个对力—能场概念的形成作出贡献的学者。虽然19世纪的德国物理学界普遍欣赏超距作用的概念，但德国存在着深刻的理性传统，在她的灵魂里是赞同场论的。我这里指的是莱布尼茨—康德的理性主义形而上学传统，这个哲学传统通过谢林(Schelling)与黑格尔的自然哲学在19世纪的德国知识分子中很有影响。

从实体与活力的统一性原理出发，莱布尼茨导出了作为力的密集连续统的实体概念。这种实体是一种动力学充盈，而不是渗透到整

个宇宙的物质的广延充盈，它以自然界的统一为基础，又把自身分化为自然界的各种粒子。类似地，康德写道：

> 物质的运动力的基本体系依赖于实体的存在（基本的原始力），这种实体是所有物质的运动力的基础……存在着一种普遍分布的在其占据或通过排斥力充满的空间内渗透一切的物质，在其所有部分均匀激活自身并且无止境地维持这种运动。[3]

[36] 上述思想在德国物理学中多次出现过。例如，迈耶坚持形而上学的力实体概念，它不依赖于物质但仍具有与物质一样的实在地位。为了同时描述通过空间传递的光与热，迈耶不仅需要一种物质以太来携带光波，也需要一种独立传递的、明显非物质的热力。迈耶关于充盈空间的力场观，在唯能论的新倾向中复活了；唯能论把能量看作是所有实在的基础，风行于19世纪的最后十年。

在德国，对场论发展的另一个刺激来自高斯（Gauss）的推测（1845），即必定存在着一些力，能够引起以有限速度在电荷之间传播的电作用。他的学生黎曼（Bernhard Riemann），早于诺伊曼就尝试发展势在时间中传播的数学表述。一开始，黎曼（1853）试图根据均匀以太对容积（引力）与形状（光）的变化的阻抗来分析引力与光的过程。他作出了这样的推理：引力可能由不可称重的、充满空间的以太进入可称重原子的连续流所构成，这依赖于直接包围可称重原子的以太的压强，而压强依次又依赖于以太的速度。因此，超距作用是不应该存在的。黎曼推测，同样的以太也应当起到传播被设想为光与热的振动的作用。然而，黎曼的推测面临着一个严重问题，它涉及新以太能够被连续创造的可能方式。

在麦克斯韦建立电动力学之前几年，黎曼更精致的统一以太场论成型了。 黎曼假设，在任何一点物体运动与运动变化的原因，应当在“连续扩散到整个无限空间的实体的运动形式”中去寻找(1858)。他把这种充满空间的实体称为以太，并且提出一种能够再现引力与光传播的偏微分方程的均匀以太的特定运动。 第一个方程是有关以太流的连续性方程。 第二个方程是关于以光速传播的速度为 W 的横波振动(不是位移的振动)的波动方程。 引力与光的这两种运动的组合产生了合乎要求的速度函数，证实了统一这两种过程的可能性。

在 1858 年提交给哥廷根科学学会的论文中，黎曼提出了光的电磁理论：

> 我已经发现，如果假设处于静止状态的电质量的作用不是瞬时发生的，而是以恒定速度(在可观测的误差极限内等于光速 c)在它们之间传播，那么电流的电动力学效应就可以得到解释。根据这个假设，电力传播的微分方程将与光和热辐射的传播方程相同。 [37]
>
> (去世后于 1867 年出版)

借助“以太中的密度变化”这个概念，黎曼建议用方程 $\nabla^2 V - 1/c^2 \partial^2 V/\partial t^2 + 4\pi\rho = 0$ 取代静电势的泊松方程 $\nabla^2 V + 4\pi\rho = 0$，而根据前一个方程，势的变化将以光速 c 从电荷向外传播(1861a)。 从而他发现，把电学和电磁学整合到他的统一物理科学纲领中是可能的。

在表征黎曼的统一纲领时，我们发现黎曼所探寻的是把超距作用的离散力代以在以太中邻近元素之间的连续推迟作用，以及以太中的状态与过程。 采用这种方式，他为力提供了本体论基础，并使得以

太的状态及其变化成为力的实在。这对能量概念来说是相当重要的转变，因为在力作为以太的状态变化穿过空间扩散时，在这个框架里力能够作为势能来理解。黎曼的纲领也是有动力学特征的。他坚持“以太运动的定律必须是为现象的解释而假设的”，而且以太动力学的基础是运动的连续性原理与势的极值条件(1858)。这些思想与威廉·汤姆孙、麦克斯韦，以及后来的菲茨杰拉德(Fitzgerald)、洛奇(Lodge)与拉莫尔(Larmor)的动力学理论的思想相似。

黎曼把以太视为空间中扩散力的一种方式的思想，把力变成了以太的能态，从而加强了把能量作为独立的守恒的实际存在物的更普遍认识，这一点正如诺伊曼的工作所表明的那样。为了描述韦伯的电以太粒子之间的表观超距力，诺伊曼以新的方式发展了黎曼早期关于传播力与推迟势的理论。在提交给哥廷根科学学会的论文(1868)中，诺伊曼认为“势[能]作为原初的、特有的动力，而力设想为次生的、动力表现自身的形式”。然后，他证明了以光速传播的势，使所有已知的电磁作用定律成为必要。以这样的方式，他使得韦伯描述光的传播的以太成为多余。虽然他可能考虑到传播势的物质基础

[38] (诸如以太场)，但是他没有涉及这一点，并因此削弱了力—能场这类物质基础的必要性。

通过忽视不可观测以太，把势能看作基本的、并可在空间中传播的东西，诺伊曼给予它一个能量场的独立地位，某种意义上类似于迈耶关于充盈空间的力场。因此，后来的唯能论者有充分的理由把诺伊曼和迈耶追认为唯能论观点的先驱。

现在我们看到，争议的焦点在于麦克斯韦与诺伊曼之间的分歧：是否存在一种本体论上不可还原的不同于物质实体的物理实体，并承担连续传递电磁作用的功能？麦克斯韦坚持实体的机械概念，在他

对诺伊曼的批评中，否定了这种可能性，因此在对电磁场的本质的新理解方面就没有贡献。

麦克斯韦，特别是威廉·汤姆孙作为普遍基质(substratum)的以太机械观在英国物理学界思维中的统治地位，不久就导致了机械世界观的兴起。在这以前，存在着两类承载具体力的电粒子，两类承载具体力的磁粒子，以及热粒子与光粒子。然而，作为威廉·汤姆孙与麦克斯韦工作的结果，这类粒子与力不久就在文献中消失了，被以太中的各类运动所取代：电磁现象被解释为以太的运动与应变；热也被视为不过是以太的运动模式。根本的假设是，所有自然现象应当可以通过普遍的以太基质的动力学得到解释。

但是，具有讽刺意义的是，追随麦克斯韦 1864 年的电磁场动力学理论的英国物理学后来的发展，能够被刻画为去机械化的、与德国特别是莱布尼茨的非机械原始实体的思想相融合的过程。例如，克利福德(W. K. Clifford)受黎曼影响。赫兹(Heinrich Hertz)在发现电磁波以后在英国也很有影响，电磁波使得赫兹与一般公众都相信场的实在性。此外，莱布尼茨的原始物质思想几乎能够在每一位英国物理学家的著作中发现，包括威廉·汤姆孙与拉莫尔的著作。

威廉·汤姆孙的普遍以太的动力学纲领认为，整个物理世界是从以太的涡旋运动中涌现出来的。作为回应，洛奇说道：

> 一种充满所有空间的连续实体：能够作为光振动；能够被剥离成正电与负电；在旋转中构成物质；连续而非碰撞地传递(物质能胜任的)每一个作用与反作用。这就是以太及其功能的现代观点。 [39]
>
> (1883)

但是，洛奇坚持“以太必定有一些不同于普通物质的性质”。因此，他暗示了以太的非机械本质。

菲茨杰拉德与19世纪80至90年代的以太理论很有关系，他也坚持“自然界的一切都可以还原为普遍充盈的运动”的观点(1885)。根据菲茨杰拉德的观点，存在着两类涡旋运动。第一是存在着构成物质的局域涡旋原子。第二是存在着通过并射向普遍流体的涡旋细线，可以赋予这种流体作为整体的以太——作为光与各种力的载体——的特征性质。也就是说，物质与以太都是作为精确满足动力学假设的麦卡拉系统(MacCullagh system)的一个结果从普遍充盈中涌现出来的，正如斯托克斯指出的那样，这个假设对于任何物质动力学系统都是不可实现的。[4] 这里我们发现了非机械以太的另一条线索。

拉莫尔是最后一位伟大的英国以太理论学家。他综合了前辈的工作，对已存在的各种潮流进行了逻辑总结。这种综合是在动力学纲领的语境中实现的，动力学纲领代表着传统机械以太理论发展的最高阶段，因而为超越机械以太理论铺平了道路。这里动力学纲领意味着拉格朗日表述。在传统上，在以太理论的拉格朗日方法中，抽象的处理迟早会尝试揭示隐机制，这种隐机制要么以实在的形式，要么至少以直观模型的方式给出由拉格朗日形式体系作出抽象处理的效应。但是，拉莫尔提出：

> 具体模型通过将无形的原始介质与其他我们能够直接观测其现象的动力学系统进行对比，来对自然的东西作出阐明与解释。
>
> (1894)

于是，情况就倒转过来了(the tail wag the dog)，这就带来了观念上

的一个关键变化：鉴于抽象的拉格朗日方法此前曾被看作通向终极理论的一步，而终极理论将表明以太是一种具体的力学系统，那么现在这个抽象的理论，在拉莫尔看来已经是最后的终极理论，而具体的模型只是出于启发或教学的缘故才仍有意义。

不久，拉莫尔更进一步地放弃英国以太理论的基本信条。1894 [40] 年，他放弃作为机械物质基础的涡旋原子。1900年，他声称“物质也许很可能是以太中的一种结构，但显然以太不是物质的一种结构”，从而加入了威廉·汤姆孙的物质以太传统的团队。“主要是在拉莫尔的影响下”，惠特克(E. T. Whittaker，1951)评论道：“接近世纪末，以太开始被广泛看作一种非物质介质，不是由在绝对空间具有确定区域的可识别元素构成的。”

这种认识帮助物理学家更严肃地考虑能量场的思想。各种形式能量的转换与守恒表明，能量体现了所有自然力的统一。因此，能量场似乎是普遍基质的合适候选者。虽然能量守恒的基础仍认为在于机械能，但是能量场的非机械含义使得这个概念渗透到物理学家的意识中，成为从机械以太场概念转向独立的、非机械的电磁场概念的必要步骤。

2.4 电磁场论

在机械以太场论中，电磁场不仅是一种根据传递局部作用的连续介质来解释表观超距作用的智力策略。而且，由于电磁场不只具有与作用的传递相关联的性质(它们中最重要的是能量以及通过场的作用来传递的具有光速并且需要经历时间的性质)，因此电磁场能够被设想为物理实在的一种表征。但是，既然电磁场被英国物理学家设想为只是机械以太的一种状态，而非独立的客体，因此我把他们的理

论称为机械以太场论，而不是电磁场论。

除了连续性，以太与普通物质的差别是非常模糊的。这在威廉·汤姆孙与拉莫尔关于物质的统一以太理论中尤其如此；在这类理论中，物质粒子要么被设想为充盈中的烟雾环或者涡旋原子(正如由磁致旋光暗示的那样)，要么被设想成处在弹性以太中旋转应变的中心。既然以太和普通物质都由力学术语来描绘，那么在它们之间就不存在清晰的差别。当然，电磁场遵守麦克斯韦方程组，而物质[41]粒子遵守牛顿运动定律。不过，既然电磁场只是被想象为遵守牛顿定律的机械以太的一种状态，因此以太理论学家总是假定电磁定律能够被还原为力学定律，尽管设计令人信服的机械模型的失败，已经将还原的任务推迟到了未来。而且，既然电荷与普通物质之间的关系超出了以太场论的范围，那么普通物质与以太之间的相互作用就只能用力学的而非电磁学的术语来处理。这进一步取消了这个理论作为电磁场论的资格。

由从麦克斯韦到拉莫尔的英国物理学家和从赫兹到洛伦兹的欧洲大陆物理学家发展起来的电磁学的动力学方法(其中对以太的力学细节未作要求)，对于电磁场论的兴起是重要的。例如，赫兹仍然承认机械以太的思想，以太能够用隐质量的运动来表示，其各部分是通过机械结构来连接的。但是，自从赫兹把麦克斯韦方程组从任何具体的力学解释中分离出来，并因而把数学形式体系完全从其机械模型的表示中分离开来，他的公理化方法培育了电磁场没有机械性质的思想的发展。

洛伦兹迈出了发展电磁场论的关键一步。与那些热衷于基于电磁场的以太的机械构造的英国物理学家相比，洛伦兹关注的是物体的电构造以及被视为对不带电物质完全透明的以太的电构造。他首先

把以太完全从物质中分离出来，而后采用他的电子论处理以太与物质之间的相互作用。在他的二元本体论中，物质体被想象成嵌入到以太中的带电粒子(离子或电子)系统，而诸如电荷甚至质量之类的性质，本质上是电磁的而非机械的。他的以太摒弃了所有机械的性质，因而完全从物质中分离出来。在这个框架中，电磁场被当作是以太的状态。既然以太没有机械性质，其性质只是如同构成电磁场与物质之基础的虚空一样，那么电磁场就享有与物质一样的本体论地位。即，它代表着独立于物质的一种物理实在，而不是物质的一种状态，像物质一样具有能量，因而具备非机械实体的资格。

洛伦兹的以太概念受若干约束，并且对于不同的解释与断言是开放的。[5] 既然这个概念对于他讨论以太与物质之间的关系是重要的，对于他关于物理世界的一般观点也是重要的，那么就值得多费一些笔墨。洛伦兹从菲涅耳的以太概念中采用了静止的特征，但作了修正。对于他来说，以太是处处局域静止的。这意味着“这种介质的每个部分都不会相对于另一部分运动，而天体的所有可感知的运动都是对于以太的相对运动”(Lorentz，1895)。因此，以太充当着给出长度与时间的正确测量的普适参考系，起着与牛顿力学中的绝对空间相同的作用。 [42]

而且，洛伦兹的以太在特殊的意义上是实体的。以太的实体性是洛伦兹的收缩假说所要求的，而收缩假说依次又是对静止以太的承认所要求的。洛伦兹假设，当物体在以太中运动时，除了电磁力，决定物体大小的分子力也受到影响，因而物体尺度的收缩最终将消除菲涅耳静止以太框架中所期望的效应。据此试图对 1887 年迈克耳孙—莫雷实验(Michelson-Morley experiment)的否定结果作出清楚解释。作为物理现象的原因，分子力的变化被设想为是由以太与物体

分子之间一些未知的物理相互作用引起的，因此以太必定是实体的，这样它才能与分子相互作用。但是，既然洛伦兹的以太对不带电的物质是完全透明的，并且没有任何关于其结构的假设，那么从物理收缩的假说中推导出来的实体性，就不足以把以太与真空区分开来。

在讨论物质与以太的关系时，洛伦兹推广了麦克斯韦的动力学方法，并假设动力学系统由两部分构成：带电粒子（电子）构成的物质和表示以太状态的电磁场。从而，两者的关系被还原为纯电磁关系，即电子与场的相互作用，不存在所假设的以太与物质间的机械联系。更精确地说，带电粒子的出现与运动，改变了以太的局域态，因而也在稍后的时间改变了这个带电粒子作用于其他带电粒子的场。也就是说，场产生于电荷的出现及其分布的变化，通过施加一种特定的“有质动”力（‘pondero-motive’ force），即所谓的洛伦兹力来作用于电荷。

[43] 洛伦兹对物质与以太关系的讨论相当于电磁场中运动电子的电动力学。在他的表述中，场的动能与势能被重新导出，并被证明是与麦克斯韦给出的结果相同的。唯一添加的是洛伦兹力的方程。因此，他的电动力学表现为麦克斯韦电磁场概念与欧洲大陆的电动力学的一种综合；在电磁场概念中，作用是以光速传播的；在电动力学中，电作用是用带电粒子之间的作用力来解释的。

洛伦兹电动力学的一个特点是，当带电粒子受到以太（通过场）的有质动力时，以太与场没有受到带电粒子的反作用。这对于坚持他的静止以太假说是必要的。庞加莱（Poincaré）批评洛伦兹，因为后者违背了牛顿第三运动定律，即作用与反作用是对等的。但是，对于洛伦兹而言，以太不是一种能够运动并且与力相互作用的质量系统。[6] 因此，对于洛伦兹来说，不仅以太与场是非机械的，而且构成

物质体的带电粒子在它们违背牛顿第三定律的意义上也是非机械的。

洛伦兹确实相信空间与时间的绝对性，而且他还坚持以太可以作为一种绝对参考系。然而，尽管如此，洛伦兹还是开始了对物理学基础的大修正，因为他拒绝牛顿运动定律的普适有效性，拒绝电磁场的机械基础，他在纯电磁的本体与概念上来设想物理学的基础。不仅电动力学方程组的有效性被设想为是基本的，无需机械解释垫底，而且自然律被还原为电磁方程组所确定的性质。特别是，力学定律被视为普适的电磁定律的特殊情况。根据洛伦兹，甚至物质体的惯性与质量，也必须由电磁术语来定义，而不能设想为恒定的。因而，牛顿力学的基本原理遭到了拒斥。

洛伦兹的电动力学在19世纪90年代被广为接受。在一个反机械论的文化氛围（马赫对力学的哲学批评和唯能论的兴起证明了这一点）中，洛伦兹的工作在电磁世界观对机械世界观的文化交锋中很有影响，并起了决定性的作用。爱因斯坦把完成始于法拉第、麦克斯韦与洛伦兹的场论纲领作为自己的使命，他晚年曾经对洛伦兹在发展场论过程中走的这一步的重要性作出评论：

> 这是令人惊奇与大胆的一步，要是没有它，以后[场论纲领]的发展将是不可能的。 [44]
>
> （Einstein，1949）

注释

1. 关于细节，见 Cantor and Hodge(1981)，Doran(1975)，Harman(1982a，b)，Hesse(1961)以及 Moyer(1978)。
2. 见 McGuire(1974)。
3. 参阅 Wise(1981)，我已经从中引用。

4. 见 Stein(1981)。

5. 对于不同的解释与断言，例如，见 Nersessian(1984)。

6. 正如斯泰因所指出的那样，洛伦兹力不是从定义动量流的麦克斯韦应力中推导出来的，否则就不会违反相当于动量守恒的牛顿第三定律。在那种情形中，因为部分动量流通过洛伦兹力把一些动量传递给运动的物体，所以场的总动量就必然改变。自 19 世纪中叶以来，能量的转化与守恒被广泛接受，这隐含了对能量的非运动学形式的接受。而这使得动量的非运动学形式合乎情理，这归因于以一定速度运动的没有确定质量的作为其载体的电磁场。因此力学与电动力学之间的表面冲突并非不可调解。详见 Stein(1981)。

第一篇

基本相互作用的几何纲领

经典场论的兴起，在寻求表观超距作用的动力因(efficient cause)方面有很深的渊源。就电磁学来说，威廉·汤姆孙与麦克斯韦在他们的以太场论中，通过引入新的实体（电磁场）与新的本体（连续以太），成功地解释了超距作用。场具有能量，因而代表物理实在。但是，作为机械以太的一种状态，场不能独立存在。在洛伦兹的电动力学中，场仍然是以太的一种状态。但是，既然洛伦兹的以太被剥夺了所有物质属性，变成了虚空(void space)的同义词，那么场就享有独立的本体论地位，与物质平起平坐。因此，在物理学研究中出现了新的研究纲领，一种建立在场本体基础上的场论纲领，与建立在粒子本体(与空间和力一起)基础上的力学纲领形成对照。 [45]

新的纲领在爱因斯坦的狭义相对论中获得了崭新的形象，其中洛伦兹以太的冗余本体从理论结构中消除了。但是，在某种意义上，场论纲领仍然没有完成。在洛伦兹的电动力学以及狭义相对论中，场必须为空间(或时空)所支撑。因此，这些场论的终极本体似乎不是场，而是空间(或时空)，或者更精确地说，是时空中的点。正如我们将在第4章中看到的那样，这种涉及场论的终极本体的隐蔽假设不是没有结果的。然而，随着广义相对论的发展，特别是运用点重合论证(point-coincidence argument)(见4.2节)，爱因斯坦证明了时空不是独立的存在，而只是作为引力场的结构性质存在。也许出于这个理由，爱因斯坦把广义相对论看作经典场论纲领的完成。我把广义相对论看作讨论20世纪场论的出发点，因为它是场在其中负责传递力的唯一本体的第一个理论。 [46]

作为传递力的因素，引力场在广义相对论中起着时空的几何结构的功能。实际上，广义相对论在场论纲领中开启了一个亚纲领——几何纲领（geometrical programme）。沿着这条线索进一步发展，

就会试图把其他力场(诸如电磁场),甚至把物质及其性质(诸如自旋与量子效应)与时空的几何结构联系起来。这个亚纲领在某些领域是成功的,诸如相对论性天文学与宇宙学,但在其他领域是失败的。而在概念上,在这个亚纲领中的研究已经提供了规范场论的几何理解的基础,并为量子场论与广义相对论的综合奠定了基础(见11.3节)。

第3章
爱因斯坦通向引力场的道路

爱因斯坦在他性格形成时期(1895—1902)意识到了物理学基础 [47] 的深刻危机。一方面，机械论观点在解释电磁学时失败了，而且这种失败招来了像马赫那样的经验主义哲学家的批评，以及诸如奥斯特瓦尔德(Wilhelm Ostwald)、黑尔姆(Georg Helm)等现象主义物理学家的批评。这些批评对于爱因斯坦评估物理学基础产生了巨大的影响。他得出的结论是，机械论观点是没有希望的。另一方面，那些追随普朗克(Max Planck)与玻尔兹曼(Ludwig Boltzmann)的科学家，他们关注可供替代的电磁论观点，同时又反对唯能论。与马赫和奥斯特瓦尔德不同，爱因斯坦相信离散的和不可观测的原子与分子的存在，并把它们作为统计物理学的本体论基础。特别是，普朗克对黑体辐射的研究，使得爱因斯坦除了认识到机械论观点的危机之外，又认识到第二个基础性危机——热力学与电动力学的危机。因此，当时的物理学“好像地面从更深处被抽掉了，看不到一块可以用来建造高楼的坚实地基”(Einstein，1949)。

爱因斯坦对于物理学基础的反思，受那个时代的两种哲学倾向指引：休谟与马赫的批判怀疑论，以及在亥姆霍兹、赫兹、普朗克与庞加莱的著作中以多种不同方式存在着的明确的康德风格。马赫对牛顿绝对空间观念的历史及概念方面的批判，冲击了爱因斯坦对已经接受的原理的信念，为他通向广义相对论铺平了道路。但是，一种经过修正的康德观点，可能对爱因斯坦的思维方式产生了甚至更为深刻与持久的影响，而且实际上塑造了他的理论化风格。根据这种观点，理性在科学理论的建构中起着积极的作用。例如，庞加莱强调"关于原理的物理学"(1890, 1902, 1904)是与热力学的胜利以及"建构性努力"的失败联系在一起的，这使爱因斯坦确信，"只有普适的形式原理的发现，才能引向可信的结果"(Einstein, 1949)。由庞加莱与爱因斯坦发现的普适原理是相对性原理(principle of relativity)。这个原理同时应用到力学与电动力学中，因而充当着作为一个整体的物理学的坚实基础。从涉及狭义相对论的惯性参考系向一般的非惯性参考系的相对性原理的延伸，通过另一个原理——等效原理(equivalence principle)——把引力场整合到这个框架中了。而这引导爱因斯坦通向广义相对论——一种有关引力场的理论。[48]

3.1 指导思想

众所周知，爱因斯坦的批判精神是在他年轻时候通过阅读休谟与马赫的著作成熟起来的。虽然休谟的怀疑论以及他对空间与时间的关系主义分析可能渗透到青年爱因斯坦的潜意识中，但马赫的影响是更切实的。尽管马赫本人喜欢坚持说他是被忽视的，然而他的经验批判主义实际上从19世纪80年代以来就产生了巨大的影响。爱因斯坦(1916d)一度评论道："即使那些自以为是马赫的反对者的人，

他们就像吮吸母乳一样，也潜移默化地吸收了马赫的观点。”是什么氛围使得马赫的影响如此巨大呢?

首先，经验主义的复兴以及现象主义态度在职业科学家中间的传播，作为对思辨的自然哲学的反动力，创造了接受马赫的经验批判主义的气氛。例如，玻尔兹曼(1888)在给基尔霍夫(Gustav Kirchhoff)的信中指出，他“将禁止所有形而上学的概念，诸如力、运动的原因”。其次，机械论的观点在解释电磁学的时候失效了；按照机械论的观点，物理实在是由物质粒子、空间、时间与力的概念来描述的。在电磁现象领域里，在赫兹探测到电磁波的实验以后，波动理论就流行了；在此理论中，物理实在被描述为连续的波场。建立在以太概念基础上的场的机械论解释，由于迈克耳逊—莫雷实验的否定结果而没了底气。作为对这种失败的反应，在电磁论观点之外，还出现了一种更为激进的唯能论观点。唯能论者拒绝承认原子的存在，相信自然界的终极连续性；他们把热力学定律作为其基础，能量被看作是唯一的终极存在，从而提供了一种反机械论的观点。

一些唯能论者(奥斯特瓦尔德、黑尔姆以及默茨)在为他们对机械论观点的激进批判辩护时，求助于实证主义(positivism)：原子与以太这类假说实体是可以取消的，因为它们的性质是无法直接观测到的。例如，默茨(1904)采取这样一种立场，呼吁“重新考虑所有物理推理的终极原理，特别是绝对运动和相对运动的力和作用的范围”。因此不难看出，马赫的经验批判主义之所以非常有影响，正是因为经验批判主义提供了一种能对物理学新分支作现象学定向解释的认识论。 [49]

但是就爱因斯坦所关注的而言，马赫的主要力量来自他对力学的历史和概念分析。作为休谟的经验怀疑论的追随者，马赫想要通过

追踪概念的历史根源来澄清可疑概念的经验起源。根据马赫的一般知识论，诸如概念、定律与理论的科学构造，只是经济地描述经验与事实信息的工具。然而，他对科学史的研究却揭示了，一些结构尽管最初只是暂时性的，但很快会得到形而上学的认可（因为成功的应用），并在实质上变得不容置疑。这样，这些结构就会从正在进行的科学活动情景中游离开来；而在科学活动情景中，批评、修正与拒绝就是法则。这些结构或者获得逻辑必然性的地位，或者作为自明的、直觉上已知的真理被接受。马赫并不接受这样一些无可辩驳的真理。

马赫的主要著作《力学的科学》（*The Science of Mechanics*）* 1883年首次出版，以对他所谓的"绝对空间的概念怪胎"的毁灭性批判而著名。绝对空间作为一种概念怪胎是因为

> 没有人能够胜任预言与绝对空间和绝对运动有关的事物；它们是纯粹的思想事物，纯粹的心智构造，不能在经验中产生。我们的所有力学原理是……涉及物体的相对位置与相对运动的经验知识。
>
> (Mach，1883)

马赫对绝对空间的批评由四点构成：(i)空间不能独立存在；(ii)只有借助宇宙中的物质客体及其性质与关系来谈论空间，才是有意义的；(iii)运动总是相对于另一个客体的；(iv)惯性力产生于相对于一个物质参考系转动或加速的客体。这些批评显示了马赫抨击传统观念以

* 即《力学的一般批判发展史》。——译者

及反对无经验根据的先验形而上学闯入科学的勇气；也显示了马赫的经验主义是批判性重估经典物理学的强有力武器。

爱因斯坦大约在1897年读到了马赫的《力学的科学》。他回忆道，"这本书对我产生了深刻而持久的影响……这归功于它面对基本概念与基本定律时的物理取向"(1952c)，而且这本书动摇了他关于力学是"所有物理思维的终极基础"的信条(1949)。据爱因斯坦说，广义相对论来自"马赫某种类型的怀疑经验主义"[与英费尔德(Infeld)和霍夫曼(Hoffmann)的谈话，1938]，而"这个理论[广义相对论]的整个思想方向与马赫的非常一致，以至于可以非常正确地认为马赫是广义相对论的先驱"(1930b)。 [50]

尽管休谟—马赫的怀疑论把爱因斯坦从对经典物理学(力学、热力学和电动力学)原理的普适性的教条式信仰中解放出来，但在他寻求物理学的坚实基础的新概念时，主要是被康德主义(Kantianism)的修正版本所指引。说真的，随着经典场论的兴起与非欧几何的发现，很多被康德视为对于经验的构造不可或缺的先验原理(诸如欧几里得几何学的公理以及有关绝对空间与超距作用的公理)都被抛弃了；人们普遍确信，大体上看来康德的先验论，尤其是他的空间观与几何观，再也站不住脚了。

然而，从科学理论的结构与认知内容来看，康德的理性主义在如下意义上在物理学与数学的发展中存活了下来：(1)由理性制定的普适原理构成了我们的经验世界，并且能够提供一幅可以充当物理学基础的协调的、连贯的世界图景；(2)实际世界的所有经验知识必须符合这些原理。虽然普适的形式原理不再被视为是固定不变的先验原理，而是作为历史演化的原理或者仅仅是约定，但是诸如普朗克与庞加莱那样很多杰出的物理学家与数学家，仍然对于理论构造的合法性

作过强有力的争论，爱因斯坦就仔细研究了他们的著作。

众所周知，在建立狭义相对论以前，爱因斯坦阅读过庞加莱的著作《科学与假说》(*Science and Hypothesis*, 1902)，这对他曾有重大的影响。[1]一些影响是明显的，另一些影响则微妙得多。前一类影响包括：我们发现庞加莱对绝对空间与绝对时间的否定，对在不同地方的同时事件的直觉否定，对绝对运动的反对，对相对运动原理与相对性原理的论述。为了进一步找到庞加莱对爱因斯坦的更微妙的影响，我们得认真地看一看庞加莱的约定主义认识论(conventionalist epistemdogy)。[2]

对于庞加莱来说，主要的认识论问题是，既然在数学和物理学中 [51] 发生了明显的分裂性变化，那么客观知识及其连续进步是如何可能的(Poincaré, 1902)？在提出这个准康德式问题时，庞加莱考虑了非欧几何的发现、机械论的危机以及他自己的原理物理学(physics of the principles)的思想(见下文)，并且拒绝了关于(绝对)空间、(欧几里得)几何学与(牛顿)物理学的正统康德立场。对于庞加莱，科学作为一个整体在本质上是经验的，而非先验的。

但是，庞加莱也完全意识到，如果不借助于智力体系的设想，无声的事件将永远不会变成经验事实；并且正如康德(1783)断言的那样，如果没有理解，感官知觉也不可能变成科学的经验。因此，根据庞加莱，一个科学理论，除了经验性假说以外，必定还涉及基本的或本质的假说或公设。与经验性假说不一样，后者作为理论的基本语言，是约定选择的结果，不是偶然的经验发现。经验事实可以引导语言的选择或变化，但是它们在这方面并没有起决定性的作用。经验事实可以决定性地影响经验假说的命运，但只是在一种选定了的语言框架之内。

那么，什么是能够据以确定科学的连续进步的科学的客观内容呢？ 庞加莱对这个问题的回答能够概括为三个概念：原理物理学、本体论的相对主义，以及结构实在论。

根据庞加莱，与渴望去发现宇宙的终极成分以及现象背后的隐机制的中心力物理学不同，原理物理学，诸如拉格朗日(Lagrange)与哈密顿(Hamilton)的分析动力学理论以及麦克斯韦(Maxwell)电磁学理论，旨在阐述数学原理，这些数学原理能够把在两个以上竞争性理论基础上取得的经验成果系统化，能够表达普通的经验内容以及这些竞争性理论的数学结构，因此这些数学原理对不同的理论解释保持中立，容许其中的任何一种解释。

原理物理学与本体论假设的无关性受到庞加莱的赞赏，因为它正好适合于他本体论的约定主义观点。 基于几何学的历史，他既不接受绝对不变的、植根于我们内心的先验本体论，也不接受能够被经验研究发现的内在本体论。 对于庞加莱，本体论假设仅仅是隐喻，它们是相对于我们的语言的，因而是可变的。

但是，在从一个旧理论及其本体论向新理论转变的过程中，除经验定律以外，一些由数学原理与公式来表达的结构关系，如果它们反映物理实在，还可以保持为真。 庞加莱以他对物理定律的不变形式 [52] 的探求而闻名。 但是，这种追求背后的哲学动机可能更加有趣。 受到马里乌斯·索弗斯·李(Marius Sophus Lie)的影响，庞加莱把描述物体集合的结构特征的转换群的不变性，看作是几何学和物理学(作为一种准几何学)的基础。 这里，我们已经触及庞加莱认识论的核心，即结构实在论(structural realism)。 这种立场有两个特征值得我们注意。 首先，我们能够拥有物理世界的客观知识。 其次，这种知识本质上只是阐明各种关系的知识。 我们能够领悟世界的结构，但

是我们不能达到自在之物(things-in-themselves)。这就进一步为庞加莱的立场只是一种康德主义修正版本的主张作了辩护。

以这样一种观点来看1887年迈克耳逊—莫雷实验以后变得明显的物理学危机，庞加莱认识到了人们所需要的是理论的重新定向或者物理学基础的转换。但是，对于庞加莱来说，所谓物理学的进步，重要的不是隐喻(诸如以太的存在或不存在)的变化，而是反映物理世界结构的强有力的数学原理表述的变化。在这些原理中，相对性原理突出地浮现在他的脑海中，而且他对这个原理的表述作出了实质性贡献。

爱因斯坦频繁地提及马赫，相比之下，他很少提到庞加莱的名字。但是，爱因斯坦理论化的理性主义风格与庞加莱的康德主义修正版本之间显著的密切关系，逃脱不了那些熟悉爱因斯坦著作的人的视线。在这方面，一个说明性文本是爱因斯坦在牛津大学的斯宾塞演讲(Herbert Spencer Lecture)，标题为“论理论物理学的方法”(On the method of theoretical physics，1933)，其中他把自己的立场概括如下：“经验内容及其相互关系都必须在理论的结论中表示出来。”但是，理论体系的结构是“理性的作品”，“尤其是理论所根据的概念与基本原理都是……人类心智的自由发明”。“经验可以提示合适的数学概念，但是数学概念大部分肯定是不能从经验推演出来的……创造性的原理存在于数学之中……我认为纯粹思想能够把握实在这种看法是正确的”，这在基础上更为统一。作为像庞加莱那样的发生经验主义者(genetic empiricist)，爱因斯坦认识到在概念与原理创造中直觉的跳跃，必定是由可感事实的整体经验所指引的，并且“经验当然保持一种作为数学建构的物理效用的唯一判据”。但是，既然“自然是最简单的可想象的数学观念的实际体现……那么，

纯粹的数学建构……提供了理解自然现象的钥匙”。

但是，强调这种密切关系，不应当模糊我们对爱因斯坦与庞加莱 [53] 之间的理论化风格的差异的认识。主要差异存在于如下事实中：从整体上看，爱因斯坦比庞加莱更为经验主义。这在他们关于具体理论的结构的断言上尤其真实。在狭义相对论的情形中，关于以太的地位，爱因斯坦不像庞加莱那样漠不关心。爱因斯坦不仅比庞加莱更严肃地处理以太这个假说性的物理实体，而且他把多余的但逻辑上可能的以太概念取消的意愿，也反映了马赫的思维经济原理的经验主义影响。在广义相对论的情况中，涉及物理空间的几何结构时，爱因斯坦就明显偏离了庞加莱的几何约定论，而采取了物理空间的几何学的经验主义概念。也就是说，他把关于物理空间的几何学看作是由引力场所构成的与决定的(见 5.1 节)。

3.2 狭义相对论

除了爱因斯坦，狭义相对论还与洛伦兹(变换)、庞加莱(群)、闵可夫斯基(时空)的名字不可分割地联系在一起。闵可夫斯基(Minkowski)的著作(1908)是爱因斯坦理论的数学说明。爱因斯坦通过仔细研究洛伦兹(如 1895)与庞加莱(如 1902)的著作，使自己认识到那些对于狭义相对论的表述来说核心的问题与重要的思想。庞加莱的狭义相对论(1905)是洛伦兹 1904 年工作的成品，后者又是庞加莱(1900)对洛伦兹早期工作批评的结果。但是，洛伦兹在他(1892, 1895, 1899)论文中的动机之一是解决由迈克耳孙—莫雷实验的否定结果引起的困惑。虽然迈克耳孙—莫雷实验能被用作其他的论证[3]，但至少在历史上，它是最后通向狭义相对论表述的一系列发展的出发点。

对于这一系列发展来说，以太的概念是其核心。洛伦兹把菲涅耳的静止以太看作电磁场的载体，并把以太与电子都看作物理世界的终极构成(所谓的电磁世界观)。虽然以太被洛伦兹赋予一定程度的实体性，但它仍然与能量、动量无关，也与除了绝对静止以外的所有其他机械性质无关。因此，洛伦兹理论中不可动以太的实际功能只是绝对空间的体现，或者作为一个优先参考系，使得物理客体的状[54]态，不论它们是处在静止还是运动状态，都能由此确定；而且在这个参考系中，时间被视为真实的时间，光速是恒定的。

因此，与机械以太不同，洛伦兹以太允许电磁场享有与物质同等的本体论地位。但是，放弃牛顿相对空间的概念(所有惯性参考系相互之间以恒定速度做匀速直线运动时，在对力学现象的描述上是等价的，没有一个比另一个优越)，对洛伦兹的电磁世界观造成了一个大难题。借助于不可动以太的概念，恒星的光行差就能够用相对于以太运动的物体对光波的部分曳引(partial entrainment)来解释(见Whittaker，1951)，但不可动以太的概念，简直与力学中有效的相对性原理不相容，并且隐含着存在物体相对于以太作绝对运动的效应。迈克耳孙—莫雷实验表明，借助于光学(或者等价于电磁学)的手段探测不到这类效应。这暗示着相对性原理似乎在电动力学中也是有效的。如果情况是这样，以太将被剥夺它作为优先参考系的最后功能而成为多余的，虽然它与电动力学的其他方面并不矛盾。

洛伦兹在承认以太的特权地位的前提下，试图解释电动力学中的表观相对性——通过假设绝对运动的真实效应为其他物理效应所补偿，运动物体在其穿越以太运动的方向上收缩，以及在穿越以太的运动中时钟明显变慢——试图解释用光学方法探测物体相对于以太的绝对运动的效应一再失败的原因。洛伦兹认为，这些效应是由于“系

统粒子与以太之间的作用力”所引起的，并用他新提出的二阶空间变换规则来表示(Lorentz，1899)。这里的经验主义倾向与洛伦兹立场的特设性本质是明显的：在电磁过程中，观测的相对性不是被看作普遍有效的原理，而是作为一些特定假设的物理过程的结果，这些物理过程的效应互相抵消，因此原则上不可观测。

庞加莱(1902)批评了洛伦兹的特设性假说，鼓励他(i)寻求一个广义理论，其原理能够解释迈克耳孙—莫雷实验的零结果；(ii)证明不论系统处于静止还是匀速平移运动状态，在系统中发生的电磁现象能够用同样形式的方程来描写；以及(iii)发现处于相对匀速运动的参考系之间的变换规则，使得方程形式对这些变换是不变的。 [55]

洛伦兹(1904)的论文被构思成与庞加莱的相对性原理相一致，根据这个原理，“光学现象只依赖于物质体的相对运动”(Poincaré，1899)。除了一些小的错误，洛伦兹在数学上成功地提供了一组所谓的洛伦兹变换(Lorentz transformations)的变换规则；在这个变换下，电动力学方程是不变的。但是，洛伦兹自己对这组变换所作解释的精神，却与相对性原理不一致。

造成这种失败的主要原因是，洛伦兹变换以不同的方式处理不同的参考系。在坐标轴固定在以太上的系统 S 中，定义了真实的时间；而在相对于 S 运动的系统S'中，尽管把变换了的空间量看作是代表真实的物理变化，但对变换了的时间坐标(地方时)却没有赋予与真实时间同样的意义，而只是看作出于数学约定的辅助量。因此，洛伦兹的解释是不对称的；在这个意义上，不仅空间变换与时间变换具有不同的物理地位，而且两个系统之间的变换不能被看作本质上是互逆的，从而不能形成一个群，也不能保证电动力学定律的严格不变性。而且，洛伦兹把他的变换限制在电动力学定律上，并坚持力学

系统仍然遵循伽利略变换(Galilean transformations)。

出于这些理由，洛伦兹(1915)承认，“我没有严格和普遍真实地证明相对性原理”，而且正是“庞加莱……[他]得到了电动力学方程的完美不变性，并表述了‘相对性公设’(postulate of relativity)，这个术语是他第一个使用的”。

这里，洛伦兹指出了两个事实。首先，在从1895年到1904年期间，庞加莱根据以下思想表述了相对性原理：(i)采用任何物理方法探测绝对运动的不可能性；(ii)物理定律在所有惯性参考系中的不变性。[4]如果我们也注意到庞加莱的光速不变原理这一公设(1898)[5]，那么正如吉迪明(Jerzy Giedymin，1982；又见Zahar，1982)令人信服地论证的那样，相对性原理的非数学表述以及光速有限与不变原理这一公设的提出就应当归功于庞加莱。

其次，在庞加莱的论文(1905)中，他从一开始就假设“相对性原理无限制地成立，根据这个原理，地球的绝对运动无法由实验得到证实”。从这样一种把相对性理解为普适原理的观点出发，庞加莱给[56]出了精确的洛伦兹变换，通过加强群的必要条件对它们作了对称的解释，并把它们表述为变换群的形式，而在这种变换群下电动力学的定律是不变的。因此，庞加莱在数学形式上与爱因斯坦同时阐述了狭义相对论。

庞加莱的成就不是偶然的，而是长期智力探索的结果，这开始于他对绝对空间与绝对时间的摒弃，并持续地批评绝对同时性(1897，1898，1902)[6]，他对绝对同时性的批评是建立在光速不变的假设基础之上的。一般来说，这种探索的驱动力就在于他日益增长的对电磁世界观的怀疑，尤其是对以太之存在的怀疑，而这种怀疑论受他的约定主义认识论指引。在《科学与假说》中，庞加莱宣称：

我们不关心以太实际上是否存在；那是一个涉及形而上学的问题。对于我们来说，重要的事情是……[以太]只是一个约定的假说，尽管人们不会停止这样做，但是以太作为无用之物而被废弃的那一天无疑会来临。

(1902)

庞加莱对以太假说是不关心的，因为，与他的原理物理学观点相符合(见3.2节)，他的主要兴趣在于为两个或多个竞争性理论所共有的、在普遍原理的形式上是可陈述的(statable)内容。因此对他来说，重要的不是以太假说本身，而是它在理论中可能承担的功能。由于以太假说与相对性原理的冲突日益明显，因此在圣·路易斯讲演(St Louis Lecture，1904)中，庞加莱就不再建议把以太看作应用相对性原理的物理系统的一部分。

无疑，爱因斯坦为他的第一篇相对论论文曾经从庞加莱的著作中汲取了灵感与支持。但是，爱因斯坦狭义相对论的同时性表述有它的优点：它在物理推理上更为简单，能够更清楚地看到相对性原理的物理内涵，尽管庞加莱的表述在数学细节上略胜一筹(见下一节)。

爱因斯坦通过快刀斩乱麻的方式解决了以太之谜而闻名天下。在他论文(1905c)的开头，他指出：

寻找地球相对于"光介质"的任何运动的不成功尝试，暗示着电动力学现象与力学现象一样，不具有任何对应于绝对静止思想的性质。它们反而暗示着……对力学方程保持成立的所有参考系，电动力学与光学的定律同样都将是有效的。我们将把这个猜想(以后称为相对性原理)提升到公设的地位……

[57] 爱因斯坦以相同的措词增加了他的第二个公设，而没有给出明显的理由。第二个公设是：光总是在虚空中以确定的速度 c 传播，而不依赖于发光体的运动状态。第一个公设已在力学传统中牢固确立起来，而第二个公设只不过是作为光的波动理论的一个结果而被接受，其中发光以太是光在其中传播的介质。爱因斯坦在第二个公设中没有提及以太，而是谈到任何一种惯性系。既然伽利略变换不能解释光速的恒定性，那么具有表观局域时间的洛伦兹变换就不得不充当联系惯性系之间现象的角色。

在洛伦兹变换下，电动力学方程仍保持不变，但力学定律变了，因为在洛伦兹电子论中，力学定律依照伽利略变换而变换。因此，爱因斯坦发现，对于以太之谜来说，关键在于洛伦兹所使用的时间概念，洛伦兹理论担负着特设性假说、不可观测效应和不对称性。电动力学定律与力学定律的变换规则依赖于两个不同的时间概念：前者是局域时间的数学概念，后者是真实时间的物理概念。然而，爱因斯坦把相对性原理延伸到电动力学中，则隐含着与洛伦兹变换有关的惯性系的等价性。于是，局域时间就是真实时间，而真实时间在不同惯性系里是不同的，因为每一个惯性系各自依赖于它们的相对速度。这意味着不存在绝对时间与绝对同时性，时间的概念必须被相对化。

由于时间的相对化，又由于洛伦兹空间变换不再被解释为由特设的分子力所引起的物理收缩，导致爱因斯坦得出这样的结论，即没有任何一种运动状态（甚至也包括静止状态）能够被指派给以太。因此，以太作为表述电动力学定律的优先参考系被证明是多余的，实际上只不过是没有物理过程与之相联系的虚空而已。

以太的废除同时削弱了机械世界观与电磁世界观，而且给物理学

的统一提供了基础。在电动力学的情形中，爱因斯坦的两个公设是与建立在空间与时间的相对概念基础之上的，并与同洛伦兹变换相联系的新运动学和谐一致。对于力学与物理学的其他部分，建立在绝对时间与绝对同时性(虽然空间概念是相对的)基础之上的、并与伽利略变换相联系的旧运动学，则是与爱因斯坦的公设相冲突的。建立在爱因斯坦两个公设基础之上的物理学的统一，要求力学和物理学其他部分与新运动学相适应。这种适应的结果是相对论性物理学，其 [58] 中同时性依赖于运动系统的相对状态，速度变换与合成遵守新的规则，甚至质量也不再是物理实体的不变性质，而是依赖于它被描述的系统的运动状态，因而是可变的。

3.3 狭义相对论的几何解释

根据费利克斯·克莱因(Felix Klein)(见附录 A1)，n 维流形的几何学是由 n 维流形的变换群的不变性理论来表征的。克莱因简单优美的方法迅速被同时代的大多数数学家接受。这个思想一个令人惊奇的应用是庞加莱与闵可夫斯基的发现，即狭义相对论不过是四维时空流形的几何学[7]，它能够表征为洛伦兹群的不变性理论。[8]

在 1905 年的论文中，庞加莱引入了几个对正确掌握狭义相对论的几何意义至关重要的思想。他选择了长度与时间单位，使得 $c = 1$，开始了一场给相对论时空的对称性投下无比光辉的实践。他证明洛伦兹变换形成了一个李群(Lie group)。他把这个群表示为四维时空的线性变换群，它混合了时间和空间坐标，但保持了二次齐式 $s^2 = t^2 - x^2 - y^2 - z^2$ 的不变性。庞加莱进一步注意到，如果人们用复值函数(it)取代 t，这样(it，x，y，z)就是四维空间的坐标，那么洛伦兹变换只不过是这个空间绕着固定原点的转动。除此以外，

他还发现了一些物理上有意义的标量与矢量，例如电荷与电流密度，能够组合成具有洛伦兹不变性的四分量实体(后来被称为“四维矢量”)，因而为现在熟悉的相对论物理学的四维表述铺平了道路。

虽然所有这些贡献是由庞加莱作出的，然而狭义相对论的几何解释主要是由闵可夫斯基发展起来的。在1907年的论文中，闵可夫斯基理解了造成“我们关于空间与时间思想的完全变化”的相对性原理的最深刻意义：现在，“空间与时间中的世界在特定意义上是一个四维的非欧流形”。在1908年的论文中，他注意到相对性公设(关于物理定律的洛伦兹不变性)“将意味着只有空间与时间中的四维世界才是由现象给出的”。

[59] 根据闵可夫斯基，物理事件恒定地包含着地点与时间的组合。他把“在某一时间点的一个空间点，即数值 x，y，z，t 的系统，称为‘世界点’(world point)”。既然任何地方、任何时刻都存在着物质，他就集中“注意力到世界点上的物质点”。从而，他得到了一幅关于“物质点的持续轨道，世界中的一条曲线，一条世界线(world-line)”的图景。他说：“根据我的意见，物理定律可能会发现它们作为这些世界线之间相互关系的最完美表达。”出于这个理由，他宁愿把相对性公设称为“世界公设”(world postulate)，它“允许对四个坐标 x，y，z，t 作相同处理”。他宣称：“从今以后，空间本身和时间本身注定消逝为只不过是阴影，只有两者的一种组合才将保持一种独立的实在性。”

由狭义相对论引起的关于空间与时间思想的变化，包括时间的相对化，使得惯性系等价(只是作为物理事件的不同部分而进入同时事件的集合)，并使得以太成为多余。结果，体现为以太的绝对空间被惯性系的相对空间取代了，尽管非惯性系仍然是相对于一个绝对空间

来定义的。在狭义相对论的闵可夫斯基表述中，每个惯性系的相对空间与相对时间能够通过单个四维流形——闵可夫斯基称之为“世界”，现在通常称之为闵可夫斯基时空——的投影得到。

闵可夫斯基认为，狭义相对论把具有独立于动力学过程的固定运动学结构赋予了四维流形。这抓住了狭义相对论作为新运动学的本质。这种运动学结构表现为一种由闵可夫斯基度规 $g_{ij} = \mathrm{diag}(1, -1, -1, -1)$表征的编时几何学(chrono-geometry)，唯一地决定了闵可夫斯基时空的惯性结构(由自由下落系统定义，并由仿射联络表征)。[9]闵可夫斯基把物理实体表示为一个世界点上的几何客体(他把它们称为时空矢量)。由于这种表示独立于参考系，它就被唯一地决定。因此，四维流形被设想为物理事件的一个舞台，物理定律被表述为几何客体之间的关系。

采用这种方式，闵可夫斯基把物理定律的洛伦兹不变性约化为四维流形中的编时几何学问题。闵可夫斯基的编时几何学是由定义为 $s^2 = g_{ij}(x^i - y^i)(x^j - y^j)$(这里 $i, j = 0, 1, 2, 3$)的闵可夫斯基间隔的洛伦兹不变性来表征的，而这种不变性整体上把 g_{ij} 固定为 $\mathrm{diag}(1, -1, -1, -1)$。欧几里得距离总是非负的，而对应的闵可夫斯基间隔对于非零的($x^i - y^i$)能够是正的、负的甚至是零，这一事实隐含着由洛伦兹(或庞加莱)群表征的编时几何学实质上是非欧几里得几何学。 [60]

在闵可夫斯基的解释中，狭义相对论是由一个在闵可夫斯基四维时空流形中定义的运动学群(即庞加莱群)表征的时空理论，这一点变得更清楚了。庞加莱群及其不变性决定了编时几何学的结构，接着这又决定了闵可夫斯基时空的惯性结构。不过，只有不存在引力时，这最后一点才是真实的。但是，在爱因斯坦的狭义相对论工作

完成之后，引力很快就成了爱因斯坦关注的焦点(1905c)。

3.4 引力场的引入：等效原理

虽然狭义相对论废除了与作为优先参考系的以太相联系的绝对空间，但只是重建了与作为一组优先参考系的惯性参考系相联系的相对空间，或者说是把静态闵可夫斯基时空作为优先参考系。如果相对性原理的精神实质就在于为在参考系的变换之下物理定律的不变形式作出论证，即否定任何优先参考系的存在，那么，爱因斯坦就询问自己："相对性原理对于相互之间加速运动的系统是否也适用？"(1907b)。

对这个研究议程的不断追求，最后导致了广义相对论的发现。在爱因斯坦的思想中，这总是涉及相对性原理的延伸，或者优先参考系的废除。从历史上讲，也是从逻辑上讲，这种追求的出发点是探究惯性质量与引力质量之间的比例关系，这个表述后来被称为等效原理。

在检验作为同时性的相对性的结果之一的质能等价性($M = m_0 + E_0/c^2$，其中 M 是运动系统的质量，E_0 是相对于静止系统的能量)的一般意义时，爱因斯坦发现"一个物理系统的惯性质量与能量，似乎表现为同一个东西"，两者均是可变的。更重要的是，他也发现，在对质量变化进行测量时，人们总是暗中假设该测量可以用天平来实现。这就相当于假设：质能等价性，

> 不仅对于惯性质量，而且对于引力质量都为真，换句话说，惯性与引力在所有情况下都是严格成比例的……惯性质量与引力质量之间的比例性，在目前所能达到的精度内，对于所有物体都是普

[61]

适有效的；……我们必须把它作为普适有效的来接受它。

(Einstein，1907b)

在惯性质量与引力质量之间的等比例性或等价性基础上，爱因斯坦在思想实验(gedanken experiment)的帮助下，提出了一种意义更深刻的等价性。

考虑两个系统 Σ_1 与 Σ_2。 Σ_1 沿着 x 轴的方向以恒定的加速度 γ 相对于保持静止的 Σ_2 运动。 在 Σ_2 中，存在着一个均匀的引力场，其中所有物体都以加速度 $-\gamma$ 下落。 爱因斯坦写道：

就我们所知，所有涉及 Σ_1 的物理定律，都与涉及 Σ_2 的物理定律没有什么不同；这归功于所有物体在引力场中经历同样的加速度这样一个事实。所以，在我们经验的当前状态，我们没有理由去假设系统 Σ_1 与 Σ_2 在哪方面是不同的。因此，从现在开始，我们同意假设，引力场与参考系的相应加速在物理上是完全等价的。

(1907b)

爱因斯坦在(1912a)论文中把这个陈述称为等效原理(简称 EP)，在1919 年的论文中，把这个思想称为“我一生中最令人愉悦的思想”。

为了延伸相对性原理，爱因斯坦借助于等效原理，引入了作为时空的惯性几何结构的基本因子的引力场。 等效原理使匀加速运动丧失了绝对性，似乎把引力场约化到相对量的地位(见下文的评论)，从而使得惯性与引力不可分离。 然而，爱因斯坦为等效原理列出的唯一经验证据是“所有物体在引力场中经历相同的加速度”。 这个事实所支撑的只是所谓的弱等效原理(weak EP)，弱等效原理要求力学

实验在一个匀加速系统中和在一个处于匹配的均匀的引力场里的静止系统中遵循相同的过程。但爱因斯坦主张的却是所谓的强等效原理(strong EP)，强等效原理假设两个这样的系统是完全等价的，以至于不可能借助于任何一种物理实验来区分它们。在下面的讨论中，等效原理总是指强等效原理。

在概念上，等效原理将“相对性原理”延伸“到参考系的匀加速平移运动的情形中”(Einstein，1907b)。等效原理的启发性价值“就在于它能够使我们用匀加速运动参考系取代任何均匀引力场”[62] (1907b)。因此，“通过对相对于匀加速运动参考系所发生的过程的理论考虑，我们得到了关于在均匀引力场中发生的过程的信息”(1911)。正如我们在4.1节中将要看到的那样，等效原理的这种解释，在爱因斯坦借助引力场来理解(时)空的(编时)几何结构方面起了指导性的作用。

起先，通过等效原理引入的经典引力场(在它类似于牛顿的标量引力场的意义上)，是一种特殊类型的坐标效应(coordinate effect)。它是在与匀加速运动参考系相联系的相对空间中定义的，而不是在四维时空中定义的。因此，场的出现依赖于参考系的选择，并且总是能够被变换掉。而且，既然等效原理只是对闵可夫斯基时空中的匀加速系有效，那么闵可夫斯基时空的惯性结构(由闵可夫斯基度规 g_{ij} 决定的仿射联络)与通过坐标变换得到的经典引力场，就只能被看作经典引力场的四维推广——广义引力场——的一个特殊例子。

人们经常论证，既然等效原理不能为引入任意引力场而应用于非匀加速与旋转的参考系，总的来看，它在把任意加速运动相对化方面失败了，它也无法消除优先参考系，因而不能作为广义相对论的基

础。[10]等效原理只有非常有限的应用范围，这的确是真的。甚至等效原理与狭义相对论的组合，也不足以表述广义相对论，这也是真的。理由很简单。等效原理所建立的只是自由下落的非旋转参考系与狭义相对论的惯性系的局域等价性。但是，表征狭义相对论的惯性系的闵可夫斯基度规，作为闵可夫斯基时空平直性的一种反映，是固定的。所以，它不能充当广义相对论的基础，因为在广义相对论中，引力场本质上是动力学的，反映的是时空的弯曲本质。[11]

然而，等效原理在方法论上的重要性存在于如下事实中，即如果可以证明(例如，通过一个与从惯性参考系到匀加速运动参考系的变化相联系的坐标系的变换)匀加速运动参考系的惯性结构是与经典引力场不可区分的，那么我们可以想象，旋转与任意加速运动参考系(通过坐标的非线性变换可得到)的惯性结构也可能等价于广义引力场：那些对应于旋转参考系的惯性结构的场将是依赖速度的，而那些对应于任意加速运动参考系的惯性结构的场将是依赖时间的。如果 [63] 情况是这样(我们将把这个假设称为等效原理的广义版本)，那么所有类型的参考系都没有固有的运动状态，而且它们都可以变换为惯性系。在这种情形下，所有参考系都是等价的，它们之间的差异仅来自于不同的引力场。

因此，等效原理所暗示的不是引力场依赖于参考系的选择，而是提示了一种设想相对空间与引力场的区分的新方法。如果承担惯性效应的结构(仿射联络)的起源在于引力场而非相对空间，那么相对空间就不具有任何固有性质，而且它们所有的结构都必须被设想为由引力场构成。这里我们发现，推广相对性原理的内在逻辑不可避免地导致引力理论。

对广义相对论来说，等效原理的意义可以分两步建立。首先，

对爱因斯坦来说，广义相对论意味着所有高斯坐标系(以及它们适合的所有参考系)都是等价的。广义等效原理已把所有具有物理意义的结构从与参考系相联系的相对空间转移到引力场，因而参考系没有任何内在的结构，能够用任意坐标系来表示。没有一个参考系比其他参考系更优先。广义等效原理取消了任何坐标系的物理意义，因而使得它们是等价的；在这个意义上，对爱因斯坦来说，它是用来获得广义相对论目标的必备工具。其次，在一个任意坐标系中，闵可夫斯基度规 g_{ij} 成为可变的。根据等效原理的观点，时空度规可以看作引力场，尽管它只是一种特殊的引力场。闵可夫斯基度规 g_{ij} 提供了扩展到一个任意引力场的四维表述的出发点。在这个意义上，等效原理也给爱因斯坦提供了通向引力理论(广义相对论的物理核心)的入口。

注意到下面这一点是有趣的：如同在相对性原理与光速不变原理的公设中一样，爱因斯坦在其等效原理的表述以及更思辨的广义等效原理的探究中，没有给我们提供从经验证据到普适原理的逻辑桥梁。爱因斯坦的原理不是“从经验中通过‘抽象’推演出来”的，这个事实经常会使人们感到困惑。但是，正如爱因斯坦后来所承认的那样：“所有概念，甚至是最接近经验的概念，都来自于自由选择约定的逻辑观点。”(1949)从爱因斯坦对等效原理的探索所作的讨论来看，我们可以感受到一些庞加莱的约定主义是如何深刻有力地影响爱因斯坦的理论化风格的气息。爱因斯坦科学研究的巨大成就也许能鼓舞科学哲学家去探索约定主义策略的富有成果的源泉。 [64]

注释

1. 例如，见 Seelig(1954)，p.69。
2. 庞加莱的认识论立场在最近几年已经变成争论的主题，注意到这一点是很有趣的。

感兴趣的读者可以在霍尔顿(Gerald Holton，1973)和米勒(Arthur Miller，1975)与吉迪明(Jerzy Giedymin，1982)和扎哈尔(Elie Zahar，1982)之间发现不一致，前者把庞加莱这位理论物理学家看作主要是归纳法优越论者，而后者把庞加莱的约定主义立场看作是他独立发现狭义相对论的决定性因素。

3. 例如，霍尔顿(1973)提供了迈克耳孙—莫雷实验与爱因斯坦的狭义相对论表述无关的强有力论证。

4. 例如，庞加莱 1895 年断言不可能测量“可称重物质相对于以太的相对运动”。1900 年，他告诉巴黎国际物理学会议，“尽管洛伦兹不这样认为”，但他相信“更精确的观测也只能揭示物质体的相对位移”(1902)。在 1904 年的圣·路易斯演讲中，他把相对性原理表述为：“根据这个原理，物理现象的定律对于静止观测者与做匀速直线运动的观测者必定是相同的，以至于我们不会也不能以任何方式决定我们实际上是否在经历这种运动。”(1904)

5. “光具有恒定的速度……这个公设不能被经验证实……它提供了确定同时性的新规则。”(Poincaré，1898)

6. “不存在绝对时间……不仅我们没有真实时间间隔相等的直觉，而且我们甚至没有发生在不同地点两个事件的同时性的直觉。”(Poincaré，1902)

7. 在 18 世纪的一些文本中，时间被明确地看作是世界的第四维。达朗贝尔(D'Alembert)在《百科全书》(*Encyclopedie*)的论文“维数”中，建议把时间看作第四维(见 Kline，1972，*Theorie des Functions Analytiques*，p. 1029)。拉格朗日在《分析力学》(*Mechanique Analytique*，1788)与《解析函数论》(*Theorie des Functions Analytiques*，1797)中，也把时间用作第四维。拉格朗日在后一部著作中指出：“因此，我们可以把力学看作四维几何学，而把分析力学看作解析几何的延伸。”拉格朗日的著作把三个空间坐标与代表时间的第四坐标放在同一个立足点上。但是，这些涉及四维的早期研究并不打算作为适当的编时几何学研究。它们只是不再与几何联系在一起的分析工作的自然推广。这一思想在 19 世纪很少受到注意，而 n 维几何学在数学家中间变得日益时髦。

8. 扩张的(非均匀的)洛伦兹群也叫庞加莱群，它是由均匀的洛伦兹变换和平移生成的。

9. 参阅 Torretti(1983)。

10. 见 Friedman(1983)。

11. 同样的思想也可用来反对等效原理的无限小概念。关于这点的详细讨论，见 Torretti(1983)，Friedman(1983)以及 Norton(1989)。

第4章
广义相对论

[65] 狭义相对论是关于闵可夫斯基时空的运动学结构的静态理论，与此相比，广义相对论作为时空几何结构的动力学理论，本质上是关于引力场的理论。从狭义相对论转向广义相对论的第一步，正如我们在3.4节中讨论的那样，是等效原理的表述；通过等效原理，匀加速运动参考系的相对空间的惯性结构，能够由静态的均匀的引力场来表示。下一步就是把等效原理的思想应用到匀速转动的刚体系统中。从而爱因斯坦(1912a)发现，匀速转动产生的静态引力场的出现，使得欧几里得几何学失效。爱因斯坦[1913年与格罗斯曼(Grossmann)一起]以其独特的理论化风格直接推广了这个结果，得出如下结论：引力场的出现一般需要一种非欧几何，并且引力场在数学上能够用一个四维黎曼度规张量 $g_{\mu\nu}$ 来表示(4.1节)。随着 $g_{\mu\nu}$ 所满足的广义协变场方程的发现，爱因斯坦(1915a-d)完成了他对广义相对论的数学表述。

把广义相对论解释为引力的几何化是非常诱人的。但是，爱因

斯坦对广义相对论的解释与此有所不同。对于他来说，“广义相对论构成了场论纲领发展的最后一步……惯性、引力以及物体与时钟的度规行为，都被约化为一个单一的场量”(1927)。也就是说，随着广义相对论的问世，不仅时空理论被约化为引力度规场的理论，而且对于任何其他场的理论，如果它被视为是相容的，那么这些场与度规场之间的相互作用就应当被严肃地看待。这个论断是通过所有场都定义在时空上这一事实得到证实的，这一事实本质上是度规场的性质的一种表现，是动力学的，而且极有可能能够与其他动力学场相互作用。

当爱因斯坦声称时空依赖于度规场时，他所意味的远不止由度规场规定的时空结构的可描述性与可确定性。相反，在爱因斯坦通向广义协变性的推理中，他最后采取了形而上学的立场，其中时空本身的存在在本体论上是由度规场构成的(4.2 节)。因此，度规场的动力学理论也是作为一个整体的时空的动力学理论，它在一开始就被设想为度规场与物质和能量的分布的一种因果关系(4.3 节)，并且场方程的各种解能够充当各种宇宙模型的基础(4.4 节)。这里，我们发现广义相对论确实给予我们一幅世界图景。根据这个图景，引力场把时空点个体化了，而且构成了可支撑所有其他场及其动力学定律的各种时空结构。 [66]

4.1 场与几何

等效原理在概念上隐含着，或者至少启发性地提示着把相对性原理从惯性系推广到任何参考系的可能性。这种推广能够通过引入作为动力因的引力场而完成，这个动力因决定着与不同物理参考系相联系的相对空间的不同物理内容。然而，为了得到广义相对性的物理

理论，爱因斯坦首先把自己从坐标必定具有直接度规意义的观念中解放出来。其次，他必须得到唯一可以承载时空的度规性质的引力场的数学描述。只有那样，爱因斯坦才能够进一步找到正确的场方程。正如施塔赫尔(1980a)指出的那样，在这条推理链中关键的一环是爱因斯坦对于旋转刚性圆盘的思考。

广义相对论一个显著而神秘的特征是张量 $g_{\mu\nu}$ 的双重功能，它一方面作为四维黎曼时空流形的度规张量，另一方面也作为引力场的一种表示。但是，直到1913年，爱因斯坦既没有提到作为引力场的一种数学表述的度规张量，也没有提到为非平直(非欧几里得)时空引入度规张量的任何必要性，这时候他仍然只关注静态引力场，这种静态引力场是建立在借助匀加速参考系来考虑等效原理的基础之上的。

甚至在狭义相对论的语境中，如同闵可夫斯基所宣称的那样，时空作为一个整体必须被设想为一种非欧几里得流形，这一点是正确的。但是，一个与惯性系相联系的、由平行类时世界线的全等性所代表的三维相对空间，本质上仍然是欧几里得的。相对空间(在应用等效原理时，与匀加速参考系相联系)也没有提供关于它们的几何本性的任何清楚的标志：无论它们本质上是欧几里得的，还是非欧几里得的。只有当爱因斯坦试图把他对引力场的研究，从静态情形扩展到非静态的、但与时间无关的情形(最简单的一种是旋转刚性圆盘问题)中时，他才确信非欧时空结构对于描述引力场是必需的。[67]

根据狭义相对论，刚体实际上不可能存在。但是，作为一种理想化，我们仍然可以定义：(i)刚体惯性运动；(ii)刚性测量杆；(iii)一个刚体惯性运动能够施行的动力学系统。[1]假定情况是这样，爱因斯坦发现：

在匀速转动的系统中，由于洛伦兹收缩[应用于周长的测量杆发生洛伦兹收缩，而应用于半径的测量杆不发生收缩]，周长与直径之比的圆周率将不得不有别于π。

(1912a)

因此，欧几里得几何的一个基本假设在这里无效。

在广义相对论的最终版本中，爱因斯坦使用了旋转圆盘论证来表明：

欧几里得几何不能应用于(一个旋转坐标系)K'。坐标的概念……预设了欧几里得几何的有效性，因而在涉及坐标系K'的时候失效了。因此，我们同样不能引入一个由相对于K'静止的时钟显示的、在K'中符合物理要求的时间。

(1916b)

但是，根据等效原理的扩展版本，

K'也可以设想为一个静止的参考系，相对于这个参考系存在着一个引力场。从而我们达到这样的结果：引力场影响甚至决定了时空连续统的度规定律。如果理想刚体的构形定律是由几何学表述的，那么在存在引力场的情况下，该几何学就是非欧几里得的。

(1921b)

爱因斯坦所熟悉的非欧几里得几何的数学处理是二维曲面的高斯理

论。不像在欧几里得平面上，笛卡儿坐标直接表示由单位测量杆测量得到的长度；而在曲面上，人们不可能以相同的方式引入坐标，以便它们具有简单的度规意义。高斯(Gauss)首先通过引入任意的曲线坐标来克服这个困难，然后把这些坐标通过后来称为度规张量的一个量 g_{ij}，与曲面的度规性质联系起来。[68] “以相似的方式，我们将在广义相对论中引入任意坐标 x_1，x_2，x_3，x_4”(Einstein，1922)。因此，正如爱因斯坦重申的那样(例如 1922，1933，1949)，如果人们(i)承认坐标的非线性变换(正如首先由旋转圆盘的例子所暗示的，并由等效原理所要求的那样)，那就导致了欧几里得几何失效；(ii)按照高斯的方式处理非欧几里得系统，那么，坐标的直接度规意义就丧失了。

曲面的高斯理论的另一个启发性作用是其内涵涉及引力场的数学描述。“在高斯曲面理论与广义相对论之间联系的最重要的一点，在于这两个理论的概念大体上都建立在度规性质之上。”(Einstein，1922)。1912 年，爱因斯坦在不知道黎曼与里奇(Ricci)或列维-奇维塔(Levi-Civita)的工作的前提下，就已经有了将广义相对论相关的数学难题与高斯曲面理论进行类比的明确思想。之后，他从格罗斯曼那里开始熟悉这些著作，并获得了关于引力场的一种数学描述(Einstein and Grossmann，1913)。

有三点对于实现这个目标是至关重要的。首先，闵可夫斯基的四维表述给出了狭义相对论的简洁描述，这被爱因斯坦看作所寻求的广义相对论的特殊情形。其次，旋转圆盘问题隐含着在引力场存在的情况下，时空结构本质上是非欧几里得的。第三，二维高斯曲面的度规性质是由度规 g_{ij} 来描述，而不是由任意坐标来描述的。而后，正如爱因斯坦向格罗斯曼寻求帮助时所看到的那样，所需要的是高斯二维曲面理论的四维推广，或者，这可以归结为同样一件事，即

由狭义相对论的闵可夫斯基表述的平直度规张量到非平直度规张量的推广。

在广义相对论的最后版本中，爱因斯坦把引力场与度规 $g_{\mu\nu}$ 联系起来。在狭义相对论有效的闵可夫斯基时空中，$g_{\mu\nu}=\mathrm{diag}(-1,-1,-1,1)$。爱因斯坦说：

> 一个自由质点相对于这个体系做匀速直线运动。因而，如果我们借助于我们选择的任意替代物，引入新的空间坐标 x_1，x_2，x_3，x_4，那么在这个新体系中，$g_{\mu\nu}$ 将不再是常数，而成了空间与时间的函数。同时，自由质点的运动在新坐标系中显现为非匀速曲线运动，而且这一运动的定律将独立于这一运动粒子的本性。从而，我们将把这种运动解释为引力场影响下的运动。因此我们发现，引力场的出现是与 $g_{\mu\nu}$ 的时空可变性联系在一起的。 [69]
>
> (1916b)

正是基于这种考虑，爱因斯坦主张“$g_{\mu\nu}$ 描述了引力场”，并且“同时还定义了所测量空间的度规性质”(同上)。

4.2 场与时空：广义协变性

用同时定义了表征时钟与测量杆行为的编时几何学结构的度规张量 $g_{\mu\nu}$ 从数学上来描述引力(或者更精确地说，是惯性-引力)结构的成功，打开了进一步研究引力场与时空结构关系的大门。然而，在1913年爱因斯坦与格罗斯曼进行合作的时候，对这个关系的澄清尚不在爱因斯坦的研究议程当中。对于他们来说，直接任务是，既然已经发现了引力场的数学描述，那就应当为作为动力学结构的场建立

方程。 在这个追求当中的指导思想是与爱因斯坦在早期研究中的思想相同的：把相对性原理(RP)从惯性参考系推广到任意参考系，以使自然界的普适定律对所有参考系都保持为真。[2]正是这个推广相对性原理的要求，引导爱因斯坦借助于等效原理的表述引入了引力场，并且把 $g_{\mu\nu}$ 看作是引力场的数学表述。 但是，随着 $g_{\mu\nu}$ 的出现，这个要求能够用数学上更精确的方式引入：$g_{\mu\nu}$ 所满足的方程，应当相对于任意的坐标变换是协变的。 也就是说，广义相对性原理要求场方程应当是广义协变的。

关于 $g_{\mu\nu}$ 的广义协变方程，很容易通过推广引力势的牛顿方程(即泊松方程)而发现：场源是通过把牛顿质量密度相对论性地推广到二阶应力—能量张量 $T_{\mu\nu}$ 得到的；因此，它们是与度规张量及其前两阶导数相关的。 正如格罗斯曼所认识到的那样，能够从后者形成的唯一的广义协变张量是里奇张量 $R_{\mu\nu}$。 爱因斯坦与格罗斯曼发现了但又拒绝了建立在里奇张量基础上的广义协变方程，因为这些方程无法约化为静态弱引力场的牛顿极限(1913)。 取而代之的是，爱因斯坦推导出了一组满足若干明显的必要条件的场方程，但是只有在有限协变群之下，才是不变的。[3]

[70]

起先，爱因斯坦因发现广义协变方程的失败而受到了干扰。 但不久以后，他感到这是不可避免的，并提出了一个反对所有广义协变场方程的简单哲学论证。 这里提到的论证就是著名的空穴论证(hole argument)。 相当有趣的是，通过首先提出而后拒斥一开始就关注因果性或唯一性或决定论的空穴论证，爱因斯坦在广义相对论的框架里面，大大澄清了他对引力场与时空关系的理解[4]，引力场与时空的关系一开始是被度规张量 $g_{\mu\nu}$ 的双重功能搞得神秘兮兮的。

下面是爱因斯坦对空穴论证的表述：

我们考虑一个时空连续统的有限部分 Σ，其中没有任何物质过程发生。如果用于描述的量 $g_{\mu\nu}$ 作为相应于坐标系 K 的坐标 x_ν 的函数给出，那么在 Σ 中物理上发生的一切，就是完全确定的。这些函数的总和用符号 $G(x)$ 表示。

假设在 Σ 之外引入一个与 K 相一致的新坐标系 K'，而在 Σ 内，新坐标系 K' 以如下方式偏离 K，使得 $g'_{\mu\nu}$ 与 K' 相关，而且像 $g_{\mu\nu}$（以及它们的导数）一样，$g'_{\mu\nu}$ 是处处连续的。$g'_{\mu\nu}$ 的总和用符号 $G'(x')$ 表示。$G'(x')$ 与 $G(x)$ 描述相同的引力场。如果我们在函数 $g'_{\mu\nu}$ 中用坐标 x_ν 取代坐标 x'_ν，即如果我们构成 $G'(x)$，那么 $G'(x)$ 也描述对应于 K 的引力场，但不对应于实际的（即一开始给出的）引力场。

现在，如果我们假设引力场的微分方程是广义协变的，那么若相对于 K 的 $G(x)$ 满足这些方程，则（相对于 K' 的）$G'(x')$ 也满足它们。于是，相对于 K 的 $G'(x)$ 也满足它们。这样一来，相对于 K，存在着两个不同的解 $G(x)$ 与 $G'(x)$，尽管这两个解在区域的边界上一致；即引力场中所发生的一切，不能唯一地被广义协变微分方程所决定。

(1914b)

该论证中有三点是关键的。首先，新张量场 $G'(x)$（"拖曳"场），能够通过场定义所在的流形（"连续统"）上的点（或主动的，active）变换（或微分同胚，diffeomorphism），从原始场 $G(x)$ 获得或产生。这个变换是如此构想的，它使得度规场 $G(x)$ 映射到了新的度规场 $G'(x)$，而应力—能量张量映射到了自身。其次，一组场方程的广义协变性(GC)要求：如果 $G(x)$ 是场方程的一个解，那么主动变换

（或拖曳）场 $G'(x)$ 也是场方程的一个解。第三，$G(x)$ 与 $G'(x)$ 是同 [71] 一个坐标系（K）里同一个点的两个物理上不同的张量场。考虑到这几点，爱因斯坦感到，广义协变性使得从应力—能量张量出发唯一地决定度规场，成为不可能的事。也就是说，广义协变性违背了因果律，因而是不可接受的。

空穴论证藉以反对广义协变性的基础是，声称 $G(x)$ 与 $G'(x)$ 在空穴（Σ）中代表着不同的物理状况。那么，怎样为这个主张作辩护呢？其实，他并没有明确地给出任何辩护。但是，不言而喻，其中假定了涉及时空本质的至关紧要的立场（或作为时空事件的流形的点的物理同一性）。即甚至在度规场在动力学上被具体说明以前，流形的点就已经被设想为在时空中物理上可识别的事件。正是这个假设使得定义在同一个点上数学上不同的场 $G(x)$ 与 $G'(x)$，应当被看作物理上有区别的场。否则，如果作为时空事件的流形上的点的物理同一性必须由动力学度规场建立起来，那么在同一个时空点定义两个不同度规场的合法性，将是一个可争论的问题，因为它依赖于人们对物理实在的理解（见下文）。

于是，一个严肃的问题是，这个假设在物理上是否可验证。在牛顿理论的语境中，既然绝对的（非动力学的）空间与时间，赋予流形上一个虚空区域的点以空间和时间性质，从而给出这些点原先就存在的物理同一性，那么这个假设就是可验证的。在狭义相对论的语境中，惯性参考系（例如，由处于刚性与非加速运动中的测量杆与时钟构成，或者检验粒子的各种组合、光线、时钟以及其他装置）扮演着与牛顿理论中绝对空间和绝对时间类似的角色，把流形的虚空区域的点个体化，因而这个假设也是可验证的。但是，在广义相对论的语境中，不存在这样能够用来把流形的点个体化的（非动力学的）绝对结

构，从而这个假设的验证是非常可疑的。但是，为了彻底检验对空穴论证的基本假设的验证，更多的哲学反思是必要的。

爱因斯坦在克服关于场方程的有限不变性困难的斗争过程中，尤其是在他努力解释1915年11月的水星近日点进动现象时，他发展了一种有关物理实在的新思想，结果成为拒斥空穴论证的关键。根据爱因斯坦(1915e，1916a与1949)，关于空间与时间的独立物理实在，在狭义相对论与广义相对论之间存在着本质的差异。在狭义相对论中，坐标(与参考系相联系)具有直接的物理(度规)意义，因为流形的虚空区域的点以惯性参考系的方式被个体化为时空点。这意味着这[72]些点具有独立于动力学度规场的预先存在的时空个体性；或者换句话说，这意味着空间与时间具有独立于动力学度规场的物理实在性。

但是，在广义相对论中，所有物理事件必须只通过质点的运动建立起来，质点的相遇[即世界线的交叉点，或时空点重合(point-coincidences)的总体]通过动力学度规场得到描述，是唯一的实在。与依赖于参考系的选择的要求相比，点重合在所有变换下(而且没有新的变换添加进去)自然得到保存，只要注意到确定的唯一性条件就行。这些考虑已剥夺了空穴论证中参考系 K 与 K' 的任何物理实在性，而且“清除了空间与时间的物理客观性的最后残余”(Einstein，1916b)。

既然广义相对论中流形的点必定从度规场中继承它们所有的差别明显的时空性质与关系，那么只有当一个四维流形是由度规场构造时，它才能被定义为一个时空。[5]也就是说，在广义相对论中，在指定度规场或求解引力场方程以前，不存在任何时空。在这种情形下，根据爱因斯坦，把一组场方程的两个不同的解 $G(x)$ 与 $G'(x)$ 归于同一坐标系或同一个流形，这是没有意义的，也不具有任何物理内

容。出于这个理由，爱因斯坦强调：

> 在连续统的同一个区域的两个不同 g 系统（更好的提法是两个不同的引力场）的（同时）实现，按照理论的本性是不可能的……如果 $g_{\mu\nu}$ 的两个系统（或者更一般地说，应用于对世界的描述中的任何变量）是如此构造的，以至于第二个系统能够通过一个纯粹的时空变换从第一个系统得到，那么它们就是完全等价的。
>
> (Einstein, 1915e)

在逻辑上，广义相对论较之度规场（它把流形点个体化为时空点）缺少结构，这就为拖曳度规场的整个等价类与一个引力场相鉴别提供了证据。这种鉴别已经损毁了空穴论证的基本假设（即拖曳度规场是物理上不同的），并挡住了空穴论证的逻辑力量（广义协变性将违背因果律），从而在爱因斯坦通向广义协变方程的道路上消除了哲学上的障碍。更进一步，既然爱因斯坦确信物理定律（方程式）应当描述并[73]确定仅仅是物理上为真的事件（时空同时事件的总体），而且物理上为真的陈述将不会"在任何（单值的）坐标变换中失效"，那么，对广义协变性的要求就变得最为自然与绝对必要了（同上）。结果，很快就得到了广义相对论的最终描述，也就是场方程的广义协变形式，即爱因斯坦场方程（Einstein，1915d）[6]：

$$G_{\mu\nu} = -k\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right) \tag{1}$$

或者，等价地为，

$$G_{\mu\nu} - \frac{1}{2}g_{\mu\nu}G = -kT_{\mu\nu} \tag{2}$$

对于爱因斯坦来说，广义协变性是推广相对性原理的数学描述。但是，这个论断，是被克雷奇曼(Eric Kretchmann，1917)以及其他后来的评论家所拒绝的。他们论证道，既然任何时空理论能够借助于纯粹数学处理，变成为一种广义协变的形式，而这些广义协变的理论包含明显违背相对性原理的牛顿时空理论的变种，那么爱因斯坦所表述的广义协变性与相对性原理之间的联系就是一种幻觉。实际上，他们申辩道，广义协变性作为物理理论一种特定的数学表述形式的性质，失去了任何物理内容。

克雷奇曼式的论证，能够通过以下两步来反驳。首先，从上述讨论中我们清楚地看到，广义协变性，正如它在广义相对论中被爱因斯坦理解并引入的那样，是物理上深刻的论证即点重合论证(point-coincidence argument)的结果。点重合论证假设时空的物理实在是由质点的世界线的交叉点构成的，这些质点依次由动力学度规场来描述。因此，从它的发生来看，广义协变性不只是一种数学要求，而且是有关时空与引力场的本体论关系的物理假设：只有被引力场描述的物理过程，才能把构成时空的事件个体化。在这个意义上，广义协变性在物理上绝不是空洞的。

其次，我们总是可以通过引入一些额外的数学结构(诸如广义度规张量或曲率张量)，把一个原来不是广义协变的理论(诸如牛顿理论或狭义相对论)变成广义协变的形式，这是确凿无疑的。但是，为了断言修正后的理论与原来的理论具有相同的物理内容，人们必须作出如下物理假设：这些附加结构只是辅助的，不具有独立的物理意义，即它们在物理上是空洞多余的。例如，在狭义相对论的情形中，有一个限制(即曲率张量必定处处消失)必须强加上去，以便能够恢复闵可夫斯基度规(与优先惯性参考系一起)。这种限制，对于恢复广义 [74]

相对论之前的理论的物理内容是必要的，因为它使得在这些理论的表述中，表观广义协变性成为平凡的且物理上乏味的。尤其值得提醒大家注意的是，由于平凡广义协变理论的解必须满足强加在这些理论表述上的限制，因此它们本身不可能是广义协变的。

通过排除平凡广义协变性的情形，广义协变性与相对性原理的推广之间的联系，似乎是深刻而自然的：既预设了并且支持了时空的关系主义观点，又拒绝了时空具有独立于引力场的物理实在的观点。[7]

4.3 物质—时空与场：马赫原理

虽然克雷奇曼不能严格证明，在广义协变方程中表达的相对性原理，只是对于没有物理内容的数学表述的一个要求，但是他的批评的确促使爱因斯坦重新考虑广义相对论的概念基础。1918 年 3 月，除了(a)相对性原理和(b)等效原理[8]之外，爱因斯坦第一次撰写了“马赫原理”(Mach's principle)的表述，并把它列出来，作为广义相对论所依托的三个主要观点之一：

> (c) 马赫原理：G 场完全由物体的质量决定。既然质量和能量一致符合狭义相对论的结果，并且能量在形式上借助于对称能量张量($T_{\mu\nu}$)来描述，那么 G 场是由物质的能量张量限制与决定的。
>
> (Einstein, 1918a)

在脚注中，爱因斯坦指出：

> 以前我没有在原理(a)与(c)之间作出区别，这是一种混淆。我已经选择了“马赫原理”的名称，因为这个原理具有马赫所要求

的惯性必须从物体的相互作用中推导出来的广义意义。

（同上）

爱因斯坦坚持马赫原理（MP）的满足是绝对必要的，因为它表达了一种涉及物质在本体论上优先于时空的形而上学承诺。他声称，场方程(1)作为马赫原理的具体化，暗示着没有任何引力场能够脱离物质而存在。他还宣称，既然宇宙中所有质量都参与到引力场的完善中，那么马赫原理就是关于整个宇宙的时空结构的陈述（同上）。

通过强调马赫原理的重要性，爱因斯坦试图澄清广义相对论的概念基础。然而，在广义相对论中马赫原理的逻辑地位，正如我们将会简要地看到的那样，远比爱因斯坦想象的复杂。不过，爱因斯坦 [75] 的努力显示了他在发展与解释广义相对论时的马赫主义动机。实际上，早在 1913 年，爱因斯坦就已经宣告了其理论的马赫主义灵感。他假设：

物体的惯性阻力能够通过在其周围放置非加速的惯性质量而增加；并且如果那些质量分享着物体的加速度，那么这种惯性阻力的增加必定消失。

(1913b)

并且宣称：

这与马赫大胆的思想相一致，即惯性起源于所考虑的物质粒子与所有其他粒子的相互作用。

(Einstein and Grossmann, 1913)

在1913年6月25日给马赫的一封信中，爱因斯坦热情洋溢地写道：

> 您对力学基础有益的研究，将得到卓越的确证。因为有必要证明惯性起源于某种相互作用，这与您关于牛顿水桶实验(Newton pail experiment)的考虑完全符合。
>
> (1913c)

1916年马赫去世时，爱因斯坦在悼辞中甚至宣称马赫“远在半个世纪以前，离推测到广义相对论就只有咫尺之遥了”(1916d)。

正如我们在3.1节中阐明的那样，马赫的思想已经对爱因斯坦的思维方式施加了深刻的影响。在马赫的思想中最重要的是，他对牛顿绝对空间概念的批评。根据马赫的观点，牛顿绝对空间概念超越了经验范围，具有不合常规的因果作用：绝对空间影响质量，但没有什么东西能够影响绝对空间。牛顿求助于与绝对空间有关的绝对运动概念来为绝对空间概念辩护，绝对空间能够产生可观测的惯性效应，诸如旋转水桶情形中的离心力。在反对牛顿的论证中，马赫取而代之地建议：

> 牛顿的旋转水桶实验只是告诉我们，水相对于桶壁的旋转不会产生任何显著的离心力，这种力只是由相对于地球与其他天体的质量的相对旋转产生的。
>
> (Mach，1883)

这里相对于绝对空间的绝对运动概念，被相对于物理参考系的相对运动概念所取代。[9]

相对运动与物理参考系都是可观测的。 但是，为什么相对运动能够产生惯性力，仍然是神秘的。 马赫的建议，正如在上述引文之 [76] 后直接暗示的那样，认为惯性力与惯性质量，以及与物理参考系相联系的相对空间中物体的惯性行为，都是由物体与其环境之间的某种因果相互作用决定的：

> 没有人能够有把握地说，如果桶壁厚度与质量增加到几里格(league)*厚的时候，实验会产生怎样的结果。
>
> （同上）

在别的地方马赫断言，有助于产生惯性力并确定物体的惯性行为的是，“与最远质量相比较，最近质量可以忽略”（Mach，1872）。 这就带来了马赫思想的另一个特征：局域惯性系统或局域动力学是由整个宇宙物质的分布决定的。

当爱因斯坦撰写了“马赫原理”的表述，并断言空间的态完全是由有质量物体决定的时，他肯定已经牢记了马赫思想的全部复杂性：空间作为物体空间关系的总体抽象(这体现在各种参考系中)，运动的相对性，以及物体的因果相互作用对惯性的决定。 但是，马赫的因果相互作用思想，是建立在瞬时超距作用的假设之上的。 超距作用的思想解释了为什么远处质量的分布能够在局域动力学(诸如惯性力)中扮演重要的角色，但是这与爱因斯坦内心珍爱的场作用概念针锋相对。 因此，为了接受马赫的思想，爱因斯坦必须按照场论的精神来重新解释它们。 这种追求不存在特别的困难，因为爱因斯坦所设想

---

* 里格是一长度单位，1 里格相当于 4.8 千米。 ——译者

的引力场很好地充当着有质量物体之间的因果作用的介质。因此，根据爱因斯坦，在引力场的概念中，"马赫的概念得到了完全的发展"(1929b)。

在爱因斯坦对马赫原理的讨论中，存在着对马赫思想的其他修正，这些修正也在场论框架内展示了爱因斯坦思想与马赫思想之间的联系。首先，在爱因斯坦场方程中所描述的能量张量，不仅是物体的能量，而且也是电磁场的能量。其次，在爱因斯坦对马赫的惯性思想(即惯性质量、惯性力以及惯性定律起源于物体的因果相互作用)的推广中，"惯性"概念同样涉及惯性系(比如不同的参考系；在这些参考系中的自然律，包括电动力学的局域定律，采取特别简单的形式)。[77] 因此，马赫的惯性思想，不仅为等效原理提供了灵感(如果因果相互作用被看作是引力相互作用)，而且也为相对性原理的延伸提供了灵感；通过这种延伸，惯性参考系被剥夺了它们的绝对特性，而变成了由度规张量(度规张量同时描述了惯性—引力场与时空的编时几何结构)表征的相对空间的特殊情形。

正如我们上面提到的，爱因斯坦把马赫原理从相对性原理中区分出来，是对克雷奇曼的批评的回应。一开始，爱因斯坦可能已经感到，相对性原理作为惯性的相对性的马赫思想的延伸，足以表征广义相对论的物理内容。但是，在克雷奇曼关于广义协变形式中的相对性原理在物理上是空洞的批评的压力下，爱因斯坦感觉到，为了避免混淆，有必要澄清广义相对论的概念基础。而这可以通过表述一个独立的、被命名为马赫原理的原理来完成，以表征广义相对论的一些指导思想——涉及(i)惯性运动(由时空的几何结构来表示)与宇宙的物质内容之间的关系；(ii)时空结构的起源与决定；(iii)时空本身的本体论地位——而同时坚持相对性原理的启发性价值。

毫无疑问，马赫的思想鼓舞并激励着爱因斯坦对于广义相对论的工作。但是，马赫原理在广义相对论中的逻辑地位是另外一回事。这是一个非常复杂而有争议的论题。人们可以追随爱因斯坦，并且不严格地声称，相对性原理是关于加速运动与转动（以及因而物理参考系）的相对性的马赫思想的推广，等效原理产生于马赫关于惯性来源于有质量物体之间的（引力）相互作用的思想，以及爱因斯坦的场方程实现了马赫关于物质是空间（时间）度规结构的唯一真实原因的思想，其中空间（时间）度规结构作为物质关系的抽象，依次描述物体的运动。[10]但是，爱因斯坦更进一步地从广义相对论出发，推导出了一些可检验的非牛顿预测。这些预测，如果得到证实，将给予马赫关于惯性的相对性思想以强有力的支持，而这使得“整体惯性，即整个 g_{ij} 场，[是]由宇宙中的物质决定的”成为必要（Einstein，1921b）。因为这个理由，它们已经被称为马赫效应（Machian effects）。

实际上，以下几种马赫效应早在 1913 年就被爱因斯坦预见到了，这还先于他在静态引力场的尝试性理论（与格罗斯曼合作）的基础上表述马赫原理。（A）当可称重质量堆叠到一个物体附近时，该物体的惯性质量必定增加（Einstein，1913b）。（B）当附近的物体加速时，一个物体一定受到加速力（Einstein，1913a）。（C）一个旋转的空心体，必定在自身内部产生一个拖曳着惯性系的科里奥利力场（Coriolis field），以及一个离心力场（Einstein，1913c）。这些效应[特别是（C）效应，即旋转质量对惯性系的拖曳作用]被爱因斯坦以及后来的评论者作为广义相对论场方程的解具有马赫行为的判据（Einstein，1922）。但是，爱因斯坦最具马赫概念显著特征的观点在他的断言（D）中表述出来： [78]

在逻辑一致的相对论中，不可能存在相对于"空间"的惯性，而只是存在质量之间彼此相对的惯性。所以，如果我在充分远离宇宙所有其他质量的地方拥有一个质量，其惯性必定降低为零。

(1917a)

但是，断言(D)是与爱因斯坦场方程多种解的存在相矛盾的。其中第一个是施瓦氏解(Schwarzschild solution)。在爱因斯坦的论文(1915c)中，得到了真空中一个静态球对称质点的近似引力场的解；通过推广这个解，施瓦西(Schwarzschild, 1916a)借助于引力势处处连续(除了在 $R = 0$ 处)的假设(R, θ 与 ϕ 都是极坐标)，得到了对一个处在别的空的时空中的孤立质点的精确解(一颗恒星或太阳系)：

$$ds^2 = (1 - \alpha/r)c^2dt^2 - dr^2/(1 - \alpha/r) - r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{3}$$

其中，

$$r = (R^3 + \alpha^3)^{1/3}, \text{ 而 } \alpha = 2Gm/c^2$$

这个解隐含着无限远处的度规是闵可夫斯基的；因此，一个试验物体具有完全的惯性，而不论它离宇宙中唯一的质点有多远。非常明显，这个解是与断言(D)直接冲突的。

爱因斯坦很早就知道施瓦氏解，而且在1916年1月9日写信给施瓦西说："我非常喜欢你对这个问题的数学处理。"(1916b)但是，这个数学处理清楚的反马赫的含意，促使爱因斯坦对施瓦西重申其马赫立场：

在最后的分析中，根据我的理论，惯性恰好是质量之间的相互

作用，而不是除了所考察的质量以外，该质量所在的“空间”也参与其中的一种“作用”。我的理论的重要特征恰恰在于，空间并不如此拥有独立的性质。

(1916b)

为了迎接施瓦氏解提出的、通过场方程(它们规定了度规在无限远处具有闵可夫斯基特征)的合理边界条件[11]表达的对马赫概念的挑战，爱因斯坦在其经典论文(1916c)中，诉诸“没有包括在我们所考虑的系统中的遥远物质”的概念。这些未知质量能够充当无限远处闵可夫斯基度规的源，从而提供了一个闵可夫斯基边界条件的马赫解释，并作为所有惯性的主要源头。 [79]

如果马赫概念被认可，那么就对遥远质量的存在提供了一个自洽的证据。人们以非常高的精度观测到，在我们所处的宇宙部分，惯性是各向同性的。但是在紧临的附近(行星、太阳、恒星)，物质不是各向同性的。因此，如果惯性是由物质产生的，那么惯性效应的绝大部分必定来自于超越了我们有效的观测之外的各向同性分布的遥远质量。

但是，正如德西特(de Sitter)指出的那样，以这种方式拯救马赫立场的代价是非常高昂的。[12]首先，普适的闵可夫斯基边界条件是与广义相对论的精神相冲突的，因为它不是广义协变的；而各种引力势必然约化为闵可夫斯基值的思想，是绝对空间的伪装版本的一个组分。一般来说，边界条件的任何详细规定，正如爱因斯坦在1916年10月末给贝索(Besso)的信中很快认识到的那样，是与作为局域物理学的广义相对论不相容的。[13]其次，遥远质量被设想为有别于绝对空间，而绝对空间是原则上独立于参考系并且是不可观测的。但是，

正如绝对空间一样，遥远质量也是原则上独立于参考系，并且是不可观测的：为了起决定无限远处引力势的作用，这些质量必须被设想为超越任何有效的观测。 因此，与牛顿的绝对空间的观念相比，马赫纲领(Machian programme)的哲学诉求已经黯然失色。

到爱因斯坦作出断言(D)的时候，他充分意识到这些困难。 爱因斯坦不顾施瓦氏解以及与边界条件的马赫概念相联系的各种困难，坚持马赫立场，理由是因为他已经发展了完成其马赫纲领的新策略(1917a)。 这一策略导致了相对论性宇宙学的产生，相对论性宇宙学被正确地看作马赫原理最富魅力的结果。 宇宙学的考虑也给了爱因斯坦对马赫原理的新信心，这在(1918a)论文中他对马赫原理的表述以及(1922)论文中对马赫效应的深入讨论中得到了证明。 但是，在转向宇宙学以前，让我们简要回顾一下马赫效应。

爱因斯坦认为，效应(A)是广义相对论的一个结果(1922)。 但是，人们已经证明这不过是一种任意选择坐标的效应。[14]实际上，惯性[80]质量在等效原理成立的广义相对论中[15]，是物体一种内在的、不变的性质，独立于环境。 为了成功地通过与遥远质量的相互作用来表述惯性质量的产生，由新的场(不同于 $g_{\mu\nu}$ 场)产生至少一个更长程的力是必要的。 但是，这就破坏了等效原理，相当于只是广义相对论的一个修正，诸如失败的布兰斯—迪克理论(Brans-Dicke theory)。[16]

在特定的初始边界条件下，效应(B)与(C)能够从广义相对论中推导出来。[17]但是，这不能简单地解释为等效原理有效的一个证明。问题在于，在这些推论中，闵可夫斯基时空被心照不宣地假设为在无限远处是有效的，这部分是在近似方法的伪装之下。 但是，这个假设，正如上面提到的，在两个方面违背了等效原理的精神。 首先，在这个语境中，无限远处的边界条件假设了牛顿的绝对空间的作用，

而且与马赫原理相冲突，这就排除了有争议的有质量物体的影响。其次，在这个语境中物质的存在，不是作为时空总体结构的唯一来源，而只是修正了后者的其他平直结构。虽然平直结构能够看作由不可观测的遥远质量描述的，但后者在哲学上的诉求不亚于绝对空间。

马赫效应确证的失败，暗示着马赫原理既不是广义相对论的必要预设，也不是广义相对论的逻辑结果。不过，有两个额外的选择仍然有利于把马赫原理整合到广义相对论中。首先，我们能够修正广义相对论，并把它变成与马赫原理相容的形式。这方面的一个努力是求助于不可观测的遥远质量，但如同上面指出的，失败了。通过引入宇宙学常量对广义相对论所作的另一个修正，将在下一节中讨论。第二个选择是，把马赫原理看作爱因斯坦场方程的解的选择原则。也就是说，马赫原理是作为这些解的外部约束，而不是作为广义相对论的本质组成部分。但是，这只有在相对论性宇宙学的语境中才有意义。因而这也将在下一节中处理。

为什么甚至爱因斯坦本人在其晚年也不欣赏马赫原理，这是由于另外一个理由。马赫原理预设了物质体是决定时空结构、甚至时空存在的唯一独立物理实体。从 20 世纪 20 年代中期开始，当爱因斯坦开始被统一场论的思想迷住以后，他就拒绝了这个前提(见 5.3 节)。他论证道，“用来代表‘物质’的[应力—能量张量] T_{ik} 总是预设了 g_{ik}”(1954a)。爱因斯坦为此给出的理由非常简单。为了决定应力—能量张量 T_{ik}，人们有必要知道控制物质行为的定律。非引力的力的存在，要求有关这些力的定律必须从外面加到广义相对论的那些定律上去。被写成协变形式的这些定律包含着度规张量的分量。结果是，除去了时空，我们将不知道物质。根据爱因斯坦后来 [81]

的观点，应力—能量张量只是总体场的一部分而不是它的源。只有总体场才构成了物理实在的终极材料，从中物质被构造出来。因此，物质在本体论上先于场的立场被坚定而明确地拒斥。从这样一种观点出发，爱因斯坦总结了他后来关于马赫原理的立场：

> 按照我的观点，我们应该不再提及马赫原理。马赫原理超前于人们认为"可称重物体"是唯一的物理实在、而且应当在理论中避免那些不能由它们完全决定的所有元素的时代。
>
> (1954a)

对爱因斯坦关于马赫原理的立场作出一些澄清是必要的。根据我的意见，事实上存在着两个马赫原理：MP_1 与 MP_2。被马赫与早期爱因斯坦所坚持、而被后期爱因斯坦所拒绝的 MP_1，宣称可称重物体是完全决定时空的存在与结构的唯一物理实在。MP_2 是爱因斯坦在统一场论时期所坚持的，宣称时空在本体论上隶属于由总体实物场代表的物理实在，在这些场的组成部分中，我们可以发现引力场，而时空的结构完全由引力场的动力学决定。有关爱因斯坦对时空及其几何结构的观点，将在 5.1 节中更多地讨论。这里有充分的理由认为，关于时空与场的关系，爱因斯坦后来的立场在精神实质上仍然是马赫的。这种立场(MP_2)与马赫的立场(MP_1)之间的唯一区别是，爱因斯坦把场而不是可称重物质看作给出时空的存在并决定时空结构的终极本体。

在 1952 年 5 月 12 日给玻恩的信中，爱因斯坦写道：

> 即使完全不知道光线的偏转、近日点进动、光谱线的移动，引

力场方程仍是令人信服的，因为它们避免了惯性系——一种作用于任何事物但事物不作用于它的幽灵。

(1952d)

不可否认的事实是，爱因斯坦成功驱除惯性系幽灵的驱魔术，是受马赫关于惯性的相对性思想激励与指引的，虽然在其原始表述 MP_1 中，这个思想既不能看作场方程的必要预设，也不能看作场方程的逻辑结果，而且正如我们将在下一节中看到的，它也不能作为有用的选择原则而被确立。

4.4 广义相对论的相容性：相对论性宇宙学的产生

[82]

既然爱因斯坦把广义相对论的初始灵感、核心承诺以及主要成就看作在于绝对空间(还有惯性系的优先地位)的废除，那么，为了拥有一个相容的理论，他必须接受两个挑战。一个是由旋转提出的疑难；旋转的绝对性质，正如牛顿及其追随者所论证的那样，需要从绝对空间结构得到支持。另一个与场方程的边界条件相联系，这里场方程的边界条件作为绝对空间的一个伪装版本起作用。为了剥夺旋转的绝对性，爱因斯坦追随马赫，求助于一直难以捉摸的遥远质量。但是，正如前面所提到的，这个策略受到德西特的尖锐批评与坚决拒绝。德西特也不赞同爱因斯坦关于边界条件的混合了马赫语境中遥远质量的思想的论述，也不赞同爱因斯坦关于相容的广义相对论应当是什么的论述。历史上，爱因斯坦—德西特的论战是相对论性宇宙学出现的直接原因。[18]

在建立场方程以后，爱因斯坦试图决定无限远处的边界条件，这是马赫关于一个质量的惯性在充分远离宇宙中所有其他质量的情况下

趋于零的思想的唯一数学表述。他清楚地认识到边界条件的规定与“一个有限的宇宙，即一个其范围已经被自然所限定、而在其中所有惯性都是真正相对的这样一种宇宙如何存在”这一问题紧密相关。[19]

在1916年9月29日与德西特的谈话中，爱因斯坦提出了 $g_{\mu\nu}$ 的一组广义协变的边界条件：

$$\begin{matrix} 0 & 0 & 0 & \infty \\ 0 & 0 & 0 & \infty \\ 0 & 0 & 0 & \infty \\ \infty & \infty & \infty & \infty^2 \end{matrix} \tag{4}$$

德西特立刻认识到它的两个含义：它的非马赫本质以及物理世界的有限性。既然在无限远处 $g_{\mu\nu}$ 的简并场(degenerate field)代表着非马赫的、不可知的绝对时空，那么“如果我们希望拥有实际世界的完全的[相容的]四维相对性，则这个世界必定是有限的”。[20]

德西特自己关于边界条件的立场建立在他的一般认识论基础上，即“超越观测范围之外的一切推断都是靠不住的”(1917e)。因为这个理由，人们必须抑制住不去断言具有普适有效性的空间无限性的边界条件；而在所考虑区域的有限空间处的 $g_{\mu\nu}$，必须在每个个别情形下被分别给出，正如在通常情形下，分别给出时间的初始条件。[83]

爱因斯坦在其“宇宙学考察”中认识到，德西特的立场在哲学上是不容置疑的。但是，正如他立即承认的那样，对这个基本问题(即马赫问题)的完全放弃，对他来说是件困难的事情。幸运的是，当爱因斯坦写出他的论文时，他已经发现从这个困难问题中走出来的方法，这是一个贯彻其广义相对论的马赫概念的连贯方法。爱因斯坦的新策略是发展一个宇宙论模型，其边界条件的缺乏，是与作为一个

整体的宇宙这个明智概念相和谐的：

> 如果有可能把宇宙看成是一个在空间广延上有限的(闭合的)连续统，则我们完全不需要任何这样的边界条件。
>
> (1917a)

那么，通过假设物质的分布是均匀的与各向同性的，爱因斯坦成功证明了：

> 相对性的普遍公设与星体低速的事实，都是与宇宙空间有限的假说相容的。
>
> (1917a)

而这标志着相对论性宇宙学的出现，它是对物质与时空之关系的哲学反思的结果。

但是，为了贯彻这个思想，爱因斯坦必须修正他的场方程，即引入宇宙学常量 λ，以便场方程(1)被转换为

$$G_{\mu\nu} - \lambda g_{\mu\nu} = -k\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right) \qquad (5)$$

引入 λ 项的理由，甚至不是为了得到一个闭合解——正如爱因斯坦在这篇论文中认识到的那样，没有 λ 项，闭合解也是可能的，亚历山大・弗里德曼(Alexander Friedman，1922)随后对此作了富有成果的探索——而是如爱因斯坦声称的，为了得到星体低速事实所需要的准静态解。在数学上，这是第一个定量的宇宙学模型，本质上与施瓦西的内部解(1916b)相同。主要的差别仅在于：内部度规在无限远

处约化为闵可夫斯基度规，而爱因斯坦的宇宙在空间上则是闭合的。[21]

λ 项的物理意义是相当模糊的。它可以被解释为真空的局域不可观测的能量密度，或者是随着距离增加的物理排斥力(负压力)，抵[84]消了引力效应，并使得物质静态分布甚至膨胀。在任何一种情形下，λ 项显然是引力场的非马赫源项，代表着一种能够作用于一切而本身不能被作用的绝对元素。但是，借助于宇宙学常量的帮助，爱因斯坦能够通过把物质视为是不受压力的，从而预言宇宙中质量密度 ρ 与其曲率半径 R 之间无与伦比的关系

$$k\rho c^2 = 2\lambda = 2/R^2 \tag{6}$$

宇宙的确定外围的这类限定已经消除了严格处理边界条件的任何必要性，而这正是爱因斯坦曾经希求的东西。

如果马赫主义动机是对宇宙结构的唯一约束，那么 λ 项的引入既不是必要的，也不是充分的。它不是必要的，因为一个闭合解，尽管不是静态的，也能够从上面提到的初始场方程(1)中得出。它不是充分的，因为修正了的场方程仍然有非马赫的解。正如德西特(1917a, b, c, d)表明的那样，当时间也被相对论化以后，即使一个系统被剥夺了物质，仍然能够得出修正后的场方程的解，

$$ds^2 = (1 - \lambda r^2/3)dt^2 - (1 - \lambda r^2/3)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7}$$

其中曲率半径 $R = (3/\lambda)^{1/2}$。该系统由德西特系统 B 来标志，相对照之下，就由系统 A 来标志爱因斯坦的宇宙模型。

爱因斯坦对模型 B 的第一个反应，正如在德西特(1917a)后记中所记录的那样，可以理解为是否定的：

> 根据我的观点，构想没有物质的世界的可能性是不会令人满意的。$g_{\mu\nu}$场应该被物质所限制，否则它根本不应当存在。

面对这个批评，德西特推理道，既然在系统 B 中，所有普通物质（星体、星云、星系群等等）都被设想为不存在，那么爱因斯坦与其他“马赫的追随者被迫假设另外的物质的存在，即存在一种世界物质（world-matter）”（1917b）。德西特写道，甚至在系统 A 中，既然宇宙学常量非常小，正如由我们关于太阳系的摄动知识所显示的那样，假设的总体质量与不可观测的世界物质必须是“如此的巨大无比，以至于与它相比，我们所有已知的物质是可以完全忽略的”（1917a）。

爱因斯坦不接受德西特的推理。在爱因斯坦给德西特的信（1917d）中，他坚持在星体之外不存在世界物质，而系统 A 中的能量密度只是存在着的星体的均匀分布。但是，如果世界物质只是普通物质的某种理想排列，那么按照德西特（1917b）的论证，系统 B 由于其方程而将完全是空的（从马赫观点来看，这将是一个灾难），而系统 A 将是不合逻辑的。不合逻辑的理由如下。既然系统 A 中的普通物质对偏离静态平衡起主要作用，那么要补偿这种偏离，物质张量就必须加以修正，从而允许内部压力与应力的存在。只有当人们引入世界物质，这才是可能的，世界物质不像普通物质，它能够等同于一种连续流。如果情况不是这样，那么内部力将没有所期望的效应，而整个系统将如同人们设想的那样不会保持静止。德西特进一步论证道，假设的世界物质“取代牛顿理论中的绝对空间或者‘惯性系’的地位。它不过是这种惯性系的物质化”（de Sitter，1917c）。[85]

先把世界物质的精确本质搁置一边，爱因斯坦在 1917 年 8 月 8 日给德西特的一封信（1917e）中，提出了反对德西特的一个论证。既

然在模型 B 中，dt^2 受系数 $\cos^2(r/R)$ 影响[22]，那么在距离 $r=\pi R/2$ 处的所有物理现象就都没有连续性。正如在引力质点的邻域附近，时钟趋向于静止那样(引力势的分量 g_{44} 趋向于 0)，爱因斯坦论证道："德西特系统无论如何无法描述一个没有物质的世界，而只能描述一个物质完全集中在表面 $r=\pi R/2$ [德西特空间的赤道]上的世界。"(1918b)可以设想的是，尽管有上面讨论的各种困难，这个论证使得爱因斯坦确信他的马赫概念，同时帮助他阐述马赫原理。

但是，德西特在如下两个基本点上拒绝爱因斯坦的批评：(i)赤道上物质的本质，以及(ii)赤道的可达性。在爱因斯坦信件(Einstein，1917e)的页边，德西特评论道：

> 如果"物质"出现在 $r=\pi R/2$ 的地方，g_{44} 必定变为 0，那么物质的"质量"必须是多大呢？我猜想是∞！于是我们接纳了一种不是普通物质的物质……而那也还将是一种遥远质量……这是拯救马赫教条的一种物质机制(*materia ex machina*)。

在(1917c，d)中，德西特也论证了赤道不是"物理上可达的"，因为一个粒子只能在"无限的时间之后到达那里，即它根本不可能到达那里"。[23]

[86] 尽管时空的马赫概念有不同的含义，但系统 A 与系统 B 分享着共同的特征：两者都是宇宙的静态模型。在概念上，一个引力世界只能依靠制衡引力吸引效应的一些物理排斥力的帮助，才能保持其稳定性。宇宙学常量被视为是能够完成这种功能的动因。[24]在物理上，这种制衡是非常微妙的事情：如果没有任何干预的话，在极微小的扰动下，宇宙将膨胀到无限大，或者收缩到一个点。[25]在数学上，

正是俄国气象学家亚历山大・弗里德曼(1922, 1924)证明了系统 A 与B 只是具有正的但变化着的物质密度的场方程的无限多解的极限情形:它们中的一些解是膨胀的,另一些解是收缩的,这依赖于制衡过程的细节。虽然亚历山大・弗里德曼自己把他的各种模型的推论看作只不过是数学练习,但实际上他提出了演化宇宙(evolving universe)的数学模型。爱因斯坦一开始抵制演化宇宙的思想,但是亚历山大・弗里德曼的严格工作迫使爱因斯坦认识到,他的场方程既允许存在整体时空结构的动力学解,也允许存在静态解,虽然他对演化宇宙的物理意义仍持怀疑态度(Einstein, 1923c)。

在 1931 年,概念情形开始变化了。在不知道亚历山大・弗里德曼工作的情况下,比利时牧师和宇宙学家勒梅特(Georges Lemaître, 1925, 1927)在研究德西特模型时,重新发现了演化宇宙的思想,并提出了一个模型,它的过去与爱因斯坦宇宙渐近,而未来则与德西特宇宙渐近。勒梅特的著作(1927)被翻译成了英语,并在 1931 年重版。这与哈勃定律(Hubble's law, 1929)的发现——强有力地暗示着一个膨胀的宇宙——和爱丁顿(Eddington)关于爱因斯坦静态宇宙的证明(1930)一起,使得演化宇宙的概念开始得到广泛的接受。毫不奇怪的是,在 1931 年,爱因斯坦宣布了他对零宇宙学常量的非静态解的偏爱(1931),而德西特宣告了他的信念:"毫无疑问,勒梅特的理论本质上是真的,并且必须作为通向更好地理解大自然的非常真实与重要的一步被接受。"(1931)

当爱因斯坦已经通过"相当艰难曲折的道路"并在 1917 年创立了相对论性宇宙学时,他写道,他对于整个宇宙模型的构造,"提供了相对论思维方式的延伸",也提供了"广义相对性能够导致一个无矛盾体系的证明"。[26]但是,德西特挑战爱因斯坦的声名,并且借助

于他自己1917年的著作，证明了广义相对论的宇宙学模型直接与马赫教条相冲突。因此，根据德西特，由于广义相对论是一个连贯的 [87] 理论，马赫教条必须放弃。德西特在1920年11月4日给爱因斯坦的信中宣布"[引力]场本身是实在的"，而不是隶属于可称重物质(1920)。这实际上显示了爱因斯坦自己曾经一直为之努力的一个方向，这开始于他的莱登演讲(Leiden Lecture，1920a)，并且终结于他后来的统一场论观念。

人们可以设想，随着人们对演化宇宙观念(这提供了一个框架，其中场方程越来越多的非马赫解[27]已经被发现)的接受，马赫原理在连贯的广义相对论中应没有地位。但情况并非如此。一些学者论证道，虽然马赫原理是与场方程的一些解相冲突的，但是为了有一个连贯的理论，我们能够不放弃马赫原理，而把马赫原理作为选择原则来消除非马赫解。例如，惠勒(John Wheeler，1964a)提出马赫原理应当被设想为"将爱因斯坦场方程可允许的解，从物理上不可接受的解中分离出来的边界条件"。对于雷恩(D. J. Raine)来说，马赫原理作为选择原则，不仅是调节性的形而上学原理。相反，他论证道，既然马赫原理包容整个宇宙，并且面对着宇宙学的证据，那么就可以用宇宙学模型来检验它，用经验数据来证实它。人们能够论证的是，这样一种马赫原理的经验地位的研究，直接显示了我们对什么是连贯的广义相对论的理解。让我们简要看一下这个论题。

马赫原理作为选择原则能够采取很多形式。[28]最重要的一种形式涉及旋转。我们先前关于场方程边界条件的讨论，显示了广义相对论是与物质宇宙只是绝对空间中的小摄动的思想不相容的。如果绝对空间存在，则我们就可以期待宇宙有某些相对于绝对空间的旋转。

但是如果马赫原理是对的，并且不存在诸如绝对空间的东西，那么宇宙必须提供一个不旋转的标准，并且它本身不能旋转。因此，类似哥德尔宇宙(Gödel universe，1949，1952)或克尔度规(Kerr metric，1963)的旋转解，是不被允许的。在这种特殊情形下，经验证据似乎支持马赫原理，因为正如柯林斯(Collins)与霍金(Hawking)的论文(1973a，b)解释的那样，观测到的微波背景辐射的各向同性，(i)(在相对于整体质量分布的每一点上局域动力学惯性系的旋转的意义上)给宇宙的可能旋转(局域涡度)设置了一个极限；(ii)证明在宇宙历史上，宇宙绝不可能在短于膨胀时间标度的时间标度内一直旋转。根据雷恩，这“提供给我们以强有力的观测证据来支持[马赫]原理”(1975)。

雷恩更进一步论证道，马赫原理也是被其他经验材料所支持的。[88] 例如，(i)各向异性地膨胀的非旋转宇宙模型是非马赫的(因为局域剪切运动酷似相对于动力学惯性系的物质的旋转)，它们被平均哈勃膨胀的剪切的极小比例的微波观测排除；而且(ii)马赫型的罗伯逊—沃克宇宙(Robertson-Walker universe)的均匀性被观测到的零星系关联所支持，星系关联度量的是星系对(pair of galaxies)在平均间距小于平均值的成协中聚集成团的趋势。因此，根据雷恩，马赫原理的这些方面也是以很高的精度被经验证实的(1981)。

也许正是如此，马赫原理作为广义选择规则的有用性，在许多没有特殊对称性的情况中还远未建立起来，因为正如林德勒(Wolfgang Rindler)所指出的那样，“雷恩的方法很难应用，除了在具有特殊对称性的情形下”(1977)。

注释

1. 见 Stachel(1980a)。

2. 爱因斯坦经常谈到坐标系(数学表示)而不是参考系(物理系统)。 但是从语境上判断，当他谈到坐标系时，真正意味的实际上是参考系。

3. 关于这些物理要求的详细说明，见 Stachel(1980b)或 Norton(1984)。

4. 参阅 Stachel(1980b)与 Earman and Norton(1987)。

5. 或者，更精确地说，只有当流形的点通过由度规场构造出来的黎曼张量的四个不变量被个体化时。 见 Stachel(1994)及其提到的参考书。

6. 这里 $G_{\mu\nu}$ 是具有 G 收缩的里奇张量，而 $T_{\mu\nu}$ 是具有 T 收缩的应力—能量张量。 方程(2)的左边现在称为爱因斯坦张量。

7. 一些评论者(Earman and Norton，1987)试图利用空穴论证来反对时空的实体主义观点，他们把公式整合到一般的广义协变性原理(GC)中，甚至在前广义相对论的理论语境中也这样处理。 他们的论证失效了，因为正如施塔赫尔(1994)指出的那样，他们没有认识到在前广义相对论的理论中非动力学具体结构的存在，这阻碍了他们通过把所有拖曳场等同于同一个引力场来避免空穴论证。 而且，在前广义相对论的理论中，时空本质上是刚性的(任何约化为某一区域的恒等式的变换，都可以约化为任何区域的恒等式)，因而空穴论证(预设了在空穴以外约化为恒等式、但在空洞内却是不等式的变换)一开始就不能应用到这些情形中。 出于这个理由，在前广义相对论的科学理论语境中，时空的本质不能由空穴论证所决定。 关于更详细的内容，见 Stachel(1994)。

8. “惯性与重量在本质上是等同的……对称的‘基本张量’($G_{\mu\nu}$)决定了空间的度规性质，以及在其中物体的惯性行为，还有引力效应。 我们将空间的态表示为描述 G 场的基本张量。”(Einstein，1918a)

9. 在马赫的讨论中，物理参考系有时是由固定恒星来表示的，在另一些场合是由宇宙中质量的总体来表示的。

10. 例如，见 Raine(1975)。

11. 既然爱因斯坦场方程是二阶微分方程，边界条件对于方程的求解是必要的。

12. 见 de Sitter(1916a，b)。

[89] 13. 见 Speziali(1972)，p.69。

14. 见 Brans(1962)。

15. 所以在广义相对论的框架中，沿着任何粒子的世界线定义一个局域坐标系，使得对于这个坐标系，物理现象遵循狭义相对论的定律，并且不显示周围质量的分布效应，总是可能的。

16. Brans and Dicke(1961).

17. 见 Lense and Thirring(1918)；Thirring(1918，1921)。

18. 关于爱因斯坦—德西特争论的精彩总结，见 Kerszberg(1989)。

19. 爱因斯坦致贝索的信(1916 年 5 月 14 日)，见 Speziali(1972)。

20. 见 de Sitter(1916a)。 爱丁顿也清楚这一边界条件思想的非马赫本质。 10 月 13 日，正当爱丁顿在等待德西特(1916a)论文的复印件时，他给德西特写信：“当你选择相对于伽利略轴旋转的轴时，你就得到了一个不能归因于可观测物质的引力场，但这一引力场在本质上归因于边界条件的补充函数——无限远处的源或穴……这对我来说，似乎反驳了可观测现象是完全被其他可观测现象所限制的根本假设。”(Eddington，1916)

21. 目前看来，爱因斯坦的静态宇宙不是闭合的。 它具有经典的宇宙时间，从而得到了“柱形宇宙”(cylindrical universe)的名称。

22. 一些数学技巧对于从方程(7)推导出这个论断是必要的，见 Rindler(1977)。

23. 与这些评论相联系的是奇点和视界问题。 有关这些问题更多的讨论在 5.4 节中给出。

24. 参阅 Georges Lemaître(1934)。

25. 参阅 Arthur Eddington(1930)。

26. 见爱因斯坦致德西特的信(1917c)，以及致贝索的信(Speziali，1972)。

27. 粗略地说，大多数宇宙学家把所有真空解(诸如德西特空间和 Taub-NUT 空间)以及平直或渐近平直的解(诸如闵可夫斯基度规或施瓦氏解)看作非马赫的。他们也把所有均匀但各向异性膨胀的解(诸如各种比安基模型)或旋转宇宙的解(诸如哥德尔模型和克尔度规)看作非马赫的。一些人也把具有非零宇宙学常量 λ 的所有模型看作非马赫的，因为这些模型中的 λ 项必须被处理为一个非马赫源的项；但是其他人对 λ 项表示了更多的容忍。关于马赫或非马赫解的判据或表述的概要，见 Reinhardt(1973)，Raine(1975，1981)与 Ellis(1989)。

28. 关于这些形式的概要，见 Raine(1975，1981)。

第5章
几何纲领

[90] 爱因斯坦的广义相对论开创了一种描述基本相互作用的新纲领，其中用几何学术语来描述动力学。在爱因斯坦关于广义相对论的经典论文(1916c)之后，这个纲领由一系列理论完成了。这一章致力于讨论这个纲领的本体论承诺(5.2节)，重新考察了它的演化过程(5.3节)，包括一些只有在爱因斯坦去世以后才开始激发起来的对广义相对论的新理解的论题(奇点、视界与黑洞)(5.4节)，但不包括目前一些试图把广义相对论量子化的努力，这类内容将简要地在11.3节中提及。考虑到爱因斯坦的工作对几何纲领的产生与发展的巨大影响，我们似乎有理由从审视爱因斯坦时空观与几何观来开始这一章(5.1节)，爱因斯坦的时空观与几何观构成了他的几何纲领的基础。

5.1 爱因斯坦的时空观与几何观

时空几何学与动力学的相关性

一般来说，一个动力学理论，不管它是否描述基本相互作用，必

定预设动力学定律与解释的一些空间几何学。实际上，一种几何学的选择预先决定或概括了它的动力学基础，即它的因果结构与度规结构。例如，在牛顿（或狭义相对论）动力学中，欧几里得（或闵可夫斯基）的时间几何学及其仿射结构，是由作为空间（时间）的运动学结构的数学描述的运动学对称群（伽利略群或洛伦兹群）决定的，它决定或反映了作为其基本动力学定律的惯性定律。在这些理论中，运动学结构与动力学无关。于是，动力学定律在运动学对称群的变换下是不变的。这意味着运动学对称性对动力学定律的形式施加了一些限制。但是，在广义相对论中，情况不是这样的。在这些理论中，没有时空的先验运动学结构，因而也就没有运动学对称性以及施加于动力学定律形式的限制。也就是说，动力学定律在任何可想象的四维拓扑流形中是有效的，因而必须是广义协变的。因此，应当指出的是，在广义相对论中对动力学定律形式施加的广义协变性的限制，本质上不同于在非广义相对论中运动学对称性施加于动力学定律形式的限制。 [91]

时空的本质与运动的本质

人们花了很长时间才认识到，一方面是时空的本质及其几何结构，另一方面是以运动的绝对性或相对性形式出现的运动的本质，在两方面之间存在着紧密的联系。这里，绝对运动的概念相对简单：它意味着非相对的或内在的运动。但是，绝对空间（时间）的概念更复杂，它可以意味着刚性的（非动力学的）、实体的（非关系的），或自主的（非相对的）。在关于空间本质的传统争论中，所有这三种意义上的绝对主义（absolutism），作为与动态主义（dynamism）、关系主义（relationalism）及相对主义（relativism）的对立面，都涉及了。[1]但是，在当代广义相对论语境的讨论中，刚性不再是有争议的问题。不过，（更强的）实体主义或（更弱的）自主主义的绝对主义版本仍然在争

论中频繁涉及。

如果只有相对运动，那么绝对空间(时间)的概念将完全不值得辩护。 而在另一方面，如果空间(时间)是绝对的，那么必定存在某种绝对运动。 但是，这里必须要做的是，不要绕道走另外一条路。 也就是说，空间(时间)的非绝对本质未必导致绝对运动的不存在，而绝对运动并不预设一种绝对空间(时间)。 尽管有这样的逻辑关系，绝对运动的存在的确要求时空必须被赋予诸如质点与惯性参考系的结构，这些结构足够支撑运动的绝对概念。 因此，涉及这些时空结构的本体论地位的一个困难问题已经产生了。 在历史上，这种概念状况经常被绝对主义者成功地加以利用。 出于这个理由，在关系主义者中间具有一个拒绝绝对运动存在的悠久传统来作为保护性策略，即使这种策略不是逻辑上必需的。[2]在采取这种策略的关系主义者当中，我们发现有马赫与爱因斯坦。

[92] **爱因斯坦的时空观**

爱因斯坦的观点具有如下特征：拒绝牛顿关于空间与时间的绝对主义观点，受到场论发展的强烈影响，以及他最后把物理场视为本体论上构成时空及其结构的物理基础。

爱因斯坦对牛顿的理解是被马赫的解释所定型的。 根据马赫的观点，牛顿坚持空间必须视为是空的，而且诸如具有几何性质的惯性系统那样的空虚空间，必须看作是基本的、独立的空间，以便人们可以进一步描述什么东西填充在空间中。[3]根据马赫的观点，牛顿力学中的空虚空间与时间扮演着至关重要的双重角色。 首先，空虚空间与时间是物理学中发生的事情的坐标系的载体，与其有关的事件是由空间与时间坐标描述的。 其次，空虚空间与时间形成了无数个惯性系中的一个，这些惯性系可以视为是从所有可想象的参考系中区分出

来的，因为它们是那些使牛顿惯性定律有效的参考系。

这导致马赫在重构牛顿立场时的第二个观点：空间与时间必定具有与质点一样的物理实在性，以使得运动的力学定律具有意义。也就是说，“物理实在”被说成是像牛顿想象的那样，一方面是由空间与时间构成的，另一方面是由相对于独立存在的空间与时间运动着的永恒存在的质点构成的。空间与时间独立存在的思想能够以这种方式得到表达：如果不存在物质，空间与时间单独也可以作为有待发生的物理事件的一种潜在舞台而存在。

而且，马赫也注意到在牛顿的运动方程中，加速度的概念扮演着基本的角色，不能单单由质点之间的间隔随时间变化所定义。牛顿的加速度，只是相对于作为一个整体的空间来设想或定义的。马赫推理(这个推理被爱因斯坦所接受)道，这使得牛顿将空间归结为运动的一种非常确定的态(即绝对静止)成为必要，而绝对静止不是由力学现象所决定的。因此，根据马赫的观点，除了空间的物理实在性，牛顿还心照不宣地把一个新的惯性决定功能归因于空间。在牛顿力学中，空间的惯性决定效应必须是自主的，因为空间作用于质量却没有东西作用于空间。爱因斯坦称这种效应为空间的“因果绝对性”(causal absoluteness)(1927)。[4]

牛顿力学被狭义相对论取代，爱因斯坦则把狭义相对论的出现视为由法拉第、麦克斯韦与洛伦兹发展起来的场论的必然结果。电磁场论在形而上学层面上不同于力学，其理由就像爱因斯坦在论文(1917b)中所概括的那样： [93]

(a) “除了质点以外”，还出现了“一种新的物理实在，即‘场’”，以及

(b)“物体之间电和磁的相互作用不是通过瞬时超距作用力实现的，而是通过以有限的速度在空间传播的过程实现的”。

考虑了(a)与(b)两点之后，爱因斯坦在狭义相对论中声称，事件的同时性不可能是以绝对性为特征的，并宣称在相对于一个惯性系做加速运动的参考系中，支配固体的定律由于洛伦兹收缩而不符合欧几里得几何学的规则。从非欧几何出发，并考虑到广义协变性与等效原理，爱因斯坦在广义相对论中提出，支配固体的定律是与引力场紧密相连的。场论的这些发展激烈地修正了牛顿关于空间、时间与物理实在的观念，而场论的更深入发展则要求对这些概念作出新的理解。正是在这种互动中，爱因斯坦发展了他关于空间、时间与物理实在的观点。

根据爱因斯坦的观点，没有物理特征的时空概念最后被取消了，这是因为：

在分离的时空点的邻域内，时空连续统的度规性质是不同的，并为存在于所讨论的区域之外的物质共同决定。这种[度规关系的]时空可变性，或者有关“空虚空间”从物理上说既不是均匀的也不是各向同性的知识……驱使我们借助于10个函数，[即]引力势$g_{\mu\nu}$，来描述它的态……空间和空间的一部分无不存在引力势，因为引力势给出了空间的度规性质；没有引力势，空间完全不可想象。引力场的存在是与空间的存在直接联系在一起的。

(1920a)

对于爱因斯坦来说，时空也被剥夺了因果绝对性。在广义相对

论中，时空的几何性质是由度规场来构造的；根据等效原理，度规场同时也就是引力场。因此，其结构物理上依赖于动力要素($g_{\mu\nu}$)的时空，就不再是绝对的。也就是说，时空不仅制约了惯性质量的行为，而且反过来，时空以及有关它的态也被惯性质量所制约。

值得指出的是，虽然爱因斯坦的时空观本质上一直是反牛顿的，并且是场论导向的，但仍经历了微妙的变化。这种变化发生于20世纪20年代中叶，显然伴随着他始于1923年的对统一场论(UFT)的追求(5.3节)。在他的前统一场论研究时期，牛顿的空虚空间概念与空间和时间的因果绝对性的概念是被拒斥的，但是空间与时间的物理实在性仍然以某种方式得到保留： [94]

> 我们的现代宇宙观认识到了两种实在，它们在概念上是互相独立的——即使它们[借助于马赫原理]是因果相联的——即引力以太与电磁场——或者如人们所称的——空间与物质。
>
> (1920a)

与这种关于物理实在的二元论立场和空间的准绝对主义观点相反，爱因斯坦在统一场论研究时期的观点，完全是关系主义者的观点："于是，空间只不过是场的四维性"(1950b)，"'事物'的一种性质(物理对象的连续性)"，以及

> 空间的物理实在是由其分量是四个独立变量——空间与时间的坐标——的连续函数的场来表示的……这正好是表示物理实在的空间性质的一种特殊依赖性。
>
> (1950a)

他甚至进一步断言：

> 空间同依存于坐标的“充满空间的东西”相反，它不是单独存在的……如果我们设想除去了引力场（即函数 g_{ik}），那么留下来的就不是类型(1) $[ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - dx_4^2]$ 的空间，而是绝对的无。
>
> (1952a)

对爱因斯坦来说，函数 g_{ik} 不仅描述了场，而且同时也描绘了时空的度规性质。因而，

> 类型(1)的空间……不是一个没有场的空间，而是 g_{ik} 场的特殊例子，对于这一点……函数 g_{ik} 具有不依赖于坐标的值。不存在诸如虚空空间的东西，即不存在没有场的空间。时空本身并不能自称为存在，而不过是场的一个结构特性。
>
> （同上）

就引力场对于时空本体论上的优先地位而论，爱因斯坦早在1915年的时候就是一位关系主义者了，那时候他拒绝了空穴论证，重新回到了广义协变性（见4.2节）。但是对于他从物理实在的二元论到一元论观点的变化起关键作用的是他对 $g_{\mu\nu}$ 场的本质与功能的新理解。如果 $g_{\mu\nu}$ 场的功能只是在本体论上构成了时空，并在数学上指明了它的度规结构（因而还有运动），那么我们是否把时空看作是 $g_{\mu\nu}$ 场的一个绝对的或关系的表示，将只是一个没有任何本质差异的语义问题。但是，如果 $g_{\mu\nu}$ 场被看作是物理实体场的一部分，能够转换为 [95]

其他物理场和从其他物理场得到(如同统一场论所主张的那样)，那么关于物理实在的二元论观点将站不住脚；时空的完全关系主义的观点将不给任何时空的自主观点留有余地，即使它仍然为定义不与绝对时空相联系、但与由物理场的总体(而非总体物理场的一部分)构成的时空相联系的绝对运动留有余地。

1952年6月9日，爱因斯坦将其时空观概括如下：

> 我希望表明，时空不必看作可以离开物理实在的实际客体而独立存在的某种东西。物理客体不是处于空间之内，但这些客体具有空间广延性。这样，“空虚空间”的概念失去了意义。
>
> (1952b)

爱因斯坦在生命的最后几年，经常坚持上述观点的基础是场论纲领，其中物理实在，或终极本体，或在物理实在的逻辑构造中不可还原的概念元素，是由场来表征的。他坚持只有这个纲领才能使脱离物质的空间概念成为多余的，因为如果人们坚持只有可称重物体才是物理上实在的，那么拒绝空虚空间的存在将是荒谬的(1950b，1952a，1953)。

爱因斯坦的几何观

与爱因斯坦的关系主义时空观紧密联系的是他独特的几何观。这种观点既非公理演绎的，也非约定论的，但本质上是实用的，试图坚持几何与物理实在之间的直接联系。爱因斯坦强调他的观点有特殊的重要性，“因为没有它，我将不能表述相对论”，并且，“迈向广义协变方程的决定性一步肯定不会跨出”(1921a)。

根据公理演绎的观点，所假设的不是有关几何学所处理的客体的

知识或直觉，而只是在纯形式的意义上(即脱离所有直觉或经验的内容)被采纳的公理的有效性。这些公理是人类思想的自由创造。几何学的所有其他命题是这些公理的逻辑结果。“非常清楚的是，单[96]单公理几何学(axiomatic geometry)的概念体系不能作出与实在客体的行为有关的断言”，因为几何学处理的客体只是被公理所定义的，与实在客体没有必然的关系(Einstein, 1921a)。

但是，对于爱因斯坦来说，几何学“把自身的存在归因于人们感觉需要认识实在客体的行为”(1921a)，“几何思想对应于自然界中或多或少精确的客体，而这些客体无疑是那些思想产生的唯一原因”(1920a)。爱因斯坦论证道，本义大地测量的“几何学”一词，赞同这样的观点，“大地测量，必须与一些自然客体相互之间各种可能的排列方式打交道”，即与地球各部分、标度等有关(1921a)。因此，为了能够对实在客体或实用刚体的行为作出断言，“几何学必须通过将公理几何学的空洞概念框架与经验实在客体对应起来，以消除它单纯的逻辑形式特征”(1921a)，并成为“支配实用刚体之空间关系的定律的科学”(1930a)。这样一种“实用几何学”(practical geometry)能够被看作“物理学最古老的分支”(1921a)。5

根据爱因斯坦，“实用几何学”的意义存在于如下事实中：它在爱因斯坦几何体与实在的实用刚体之间建立了联系，以至于“宇宙的几何究竟是欧几里得的抑或非欧几里得的问题有了清楚的意义，而且这个问题的答案只能由经验来提供”(1921a)。

在庞加莱的几何观中，没有为实用刚体与几何体之间的联系留有余地。他指出，在更细致的考察下，在本性上实在的固体在几何行为方面不是刚性的，即它们的相对排列的可能性依赖于温度、外力等。因此，几何学与物理实在之间的直接联系显然是不存在的。从

而，根据庞加莱，几何学应用于经验必定涉及关于物理现象的假说，诸如光线的传播、测量杆的性质之类的假说，既具有抽象（约定）内容，又具有经验内容，正如在每一种物理理论中一样。当一种物理几何与观测不一致时，一致性可以通过更换不同的几何，或不同的公理体系，或者修改与之相联系的物理假说而得到恢复。[6]因而，根据庞加莱的约定论几何观，不管实在客体行为的本质是什么，保留欧几里得几何学应当是可能的与合理的。因为如果理论与经验之间的矛盾本身就已显露出来，人们总是试图改变物理定律，并保持欧几里得几何学，因为在直觉上，根据庞加莱，欧几里得几何学是组织我们经验的最简单形式。[97]

爱因斯坦在原则上不能拒绝庞加莱的一般立场。他承认根据狭义相对论，不存在本性上真正的刚体，从而刚体所指示的性质不能应用于物理实在。但是，他仍然坚持：

> 要准确地决定量具的物理状态，使得相对于别的量具来说，它的性状足以毫无歧义地允许它去代替"刚体"，那并不是困难的事。这种量具正是那些关于刚体的陈述所必须参考的。
>
> (1921a)

严格地说，爱因斯坦也认识到，测量杆与时钟将不得不被表示为基本方程的解（作为由运动着的原子构型组成的客体），而不是像以前那样作为理论上自足的实体，在理论物理中扮演独立的角色。不过，他论证道，"我们还远没有得到一种关于原子结构的理论原理的可靠知识，使我们在理论上能由基本概念构成刚体与时钟"，从这些事实看来，"在理论物理发展的目前阶段，我的信念是这些概念仍然

必须作为独立的概念使用”(1921a)。

因此，为了把编时几何学与物理实在联系起来，为了给出 $g_{\mu\nu}$ 场的编时几何学意义，爱因斯坦暂时接受了实用刚性杆与时钟的存在，而它们则是由实用编时几何学来处理的。

但是，人们对此会有严重误解，并因而认为爱因斯坦坚持一种外在的（编时）几何观，（编时）几何学在被看作是由刚性杆与时钟构成的意义上是外在的。[7]从三条理由来看，这是一种误解。首先，爱因斯坦承认广义相对论中的非线性坐标变换，这种变换使得坐标的直接度规意义消失了。因此，在高斯—黎曼的意义上，时空流形的几何学独立于坐标[8]，爱因斯坦的观点是内在的而非外在的。其次，爱因斯坦坚持时空的编时几何性质只是由引力场给出的(1920b)。因此，在时空的编时几何结构被设想为独立于被测量杆与时钟探测而存在或甚至独立于测量杆与时钟存在的意义上，爱因斯坦的观点是内在的，而非外在的。再次，爱因斯坦强调，在四维时空之外还引入两类物理事物，即一方面是测量杆与时钟，另一方面是所有其他事物(例如质点与电磁场)，在某种特定的意义上是不相容的。虽然他认为一开始就承认这种不相容性会更好，但是这种不相容性在物理学的后一个阶段，即统一场论的阶段，必须被消除(1949)。

[98]

但是，如果我们接受爱因斯坦把编时几何学看作本体论上由 $g_{\mu\nu}$ 场构成、数学上由 $g_{\mu\nu}$ 来描述的立场，那么必将导致不可忽视的复杂性。首先，存在着对场方程进行解释的复杂性。既然场方程等价于用 $g_{\mu\nu}$ 表示的缩并曲率张量，$g_{\mu\nu}$ 和以对称的方式规定了物质分布的应力—能量张量一起，刻画了（编时）几何学的特征，那么就不能声称这一项比另一项更优越。因此，场方程所提示的是度规场与物质(包括所有非度规场)之间的相互因果关系。但是，爱因斯坦早期被马赫

的思想困扰，这使得他声称度规场与时空的编时几何本身必须单方面地完全由物质决定(1918a)。这似乎有利于声称爱因斯坦坚持一种外在的编时几何观。就爱因斯坦坚持马赫原理而言，这种声称具有某种合理性。

但是，爱因斯坦从真空场方程的考察中得到的印象是，度规场不能把自己的存在归于物质；并且在从马赫原理通向统一场论的过程中，爱因斯坦逐渐认识到，正如在麦克斯韦的指引下，电磁场获得了解放并赢得了作为独立动力学实体的地位一样，度规场也应该被允许借马赫之手解除它的锁链，以拥有自身权利的参与者登上物理学舞台，使它具有自身的动力学自由度。实际上，爱因斯坦在临近生命终点时曾清楚地指出，在他的统一场论中，“代表‘物质’的 T_{ik} 总是预设了度规张量 g_{ik}”(1954a)。

其次，存在着一些看待编时几何学的物理基础的考虑。爱因斯坦坚持认为时空的物理实在是由场来表征的(1950a)；时空及其编时几何学声称，只存在作为物理场的一种结构性质(1952a)，其中引力与度规只是不同的表象；而“几何的场和其他种类的场之间的差别不是逻辑地建立起来的”(1948a)。[9]如果我们记住爱因斯坦的所有这些陈述，那么说爱因斯坦坚持一种编时几何学的内在的或甚至绝对论观点的立场看来是不牢靠的。

众所周知，爱因斯坦几何观的核心是黎曼的著名论题： [99]

> 我们必须寻找在它(即形成空间背景的实际事物)之外的度规关系的基础，其中结合力作用于它。
>
> (Riemann，1854)

在爱因斯坦看来，这里“结合力”所指的显然是引力，或者更精确地说，以引力场为介质的引力相互作用。因此，似乎有可能把爱因斯坦的几何观概括如下。

爱因斯坦把引力场看作只是引力相互作用因而也是编时几何学的本体论上的充要构成要素，不需要任何测量杆与时钟作为编时几何学的必要构成要素，就此来看，他的观点可以看成是内在的。

爱因斯坦坚持编时几何学不是本体论上不可归约的，而只是引力场的结构性质的一种表现，在这个意义上，他的观点应该被视为是一种关系论。

5.2 几何纲领：强的与弱的版本

根据其本体论承诺，基本相互作用的几何纲领(GP)，其基础是由爱因斯坦在广义相对论中建立起来的，是场论纲领的一个变种；根据其动力学描述，几何学术语扮演着独特的角色。在概念上，两个假设构成了它的出发点。首先，等效原理，假设了惯性与引力的不可分离性，因而也就假设了在构成时空的几何(即惯性的或仿射的)结构中引力场的作用(见3.4节)，以及在描述引力的动力学中几何结构的作用。其次，按照“实用几何学”的观点，几何学是与实用刚性杆和时钟的物理行为直接联系在一起的，因而时空的几何结构，作为物理现象的基础，不是先验给定的，而是由物理力决定的。基于这个理由，几何学应当被看作物理学的一个分支，而物理世界的几何性质不是一个先验的或分析的或约定的问题，而是一个经验的问题。正如我们在5.1节末尾指出的那样，这个思想在黎曼著名的1854年就职演讲中有其智力起源。

从这两个假设开始，爱因斯坦发现了引力场中的粒子轨迹与光线

具有非欧几里得流形中测地线的性质，而且引力场的出现是与非欧几里得流形的时空可变的度规系数 $g_{\mu\nu}$ 联系在一起，并由度规系数描述的。在经典场论中，场是根据依赖于在预先给定的欧几里得空间或闵可夫斯基时空中的坐标的势函数来描述的；与经典场论相比，爱因斯坦在广义相对论中把引力势直接与时空的几何性质联系在一起，以便引力的作用能够用几何学术语来表示。在这个意义上，我们可以说爱因斯坦开创了引力理论的几何化，因而可称为“几何纲领”。 [100]

几何纲领的基本思想就是这样。引力(或其他)相互作用通过诸如时空曲率的特定几何结构得以实现。几何结构影响了物质运动的测地线，并被物质的能量张量所影响。后者的影响用场方程表示，而前者的影响用运动的测地线方程表示。测地线方程一开始是独立于场方程预设的(Einstein and Grossmann，1913b)，但后来证明不过是场方程的一个结果(Einstein，Infeld，and Hoffman，1938)。

几何纲领有两个版本：强版本与弱版本。根据强版本，(i)时空的几何结构本身是物理上实在的，如同物质与电磁场的实在性一样，而且具有作用于物质的实在效应；(ii)引力相互作用被看作物质运动的时空几何效应的局域测量，作为时空曲率的一种表现，并通过测地线偏差方程得到表达。

爱因斯坦自己在早年坚持这种观点。实际上，晚至 1920 年，他仍然把引力场等同于空间(1920b)，或者更精确地说，等同于定义在流形中的几何结构。依据这种观点，爱因斯坦把时空的几何结构从给定的、刚性的、不可变的和绝对的实体变换为与物质相互作用的可变的动力学场。值得注意的是，几何纲领的强版本与断言空间实在性和活动性的牛顿立场非常一致，虽然它有不同于牛顿立场之处：把空间看作是充满物理性质的动力学实在，而不是把空间看作如同牛顿

主义所想象的那样是空虚的与刚性的。

几何纲领的弱版本拒绝时空结构的独立存在，只是把它们看作场的结构性质。爱因斯坦在他追求统一场论的晚年，坚持这种立场。但是，从逻辑上讲，弱版本不必预设统一场论。它所预设的是时空的几何结构在本体论上是由物理(引力)场构成的。

[101] 于是，几何纲领的弱版本面临的严峻问题是如何验证这个预设。这种验证很容易由爱因斯坦的实用几何观给出。根据这个观点，几何结构只能通过测量杆与时钟的行为表现自身，并且在整体上一方面被引力场之间的相互作用，另一方面被测量杆与时钟所决定。测量杆与时钟，除了它们作为时空度规探针的功能，也能影响引力场。但是场本身，根据爱因斯坦，足以构成时空的度规。在空虚的时空中，正是纯引力场是唯一的物理实在，并构成了时空的度规(闵可夫斯基的或其他的)；在非空虚的时空中，正是与物质(和/或电磁场)相互作用的引力场起着同样的构造性作用。

爱因斯坦的引力理论，正如他从广义相对论诞生之日开始到生命的最后日子为止反复指出的那样，只能被看作一种场论，其中引力作为因果过程是以物理场为介质的。

我已经把爱因斯坦的引力研究纲领表征为几何的，同时也是场论的。这两类特征的相容性是显然的。在爱因斯坦的早期与晚期的观点中，场与时空的几何作为引力相互作用的介质，都是直接地和不可分割地联系在一起的。实际上，形成几何纲领两个发展阶段的两种几何纲领观点，也可以看作场纲领(field programme)的两种观点或两个发展阶段。

观点1：时空的几何结构被视为物理上实在的，引力场可约化

为时空的几何结构。显然，这种观点是与物质和空间/场的二元论相容的。

观点2：引力场被看作一种物理实体，是根源于时空的几何结构，而时空是引力场的结构性质的表现。不难看出，这种观点为通向爱因斯坦统一场论所欣赏的一元论的道路作了准备。

从爱因斯坦发表的作品判断，我们发现从观点1到观点2的转变发生在20世纪20年代的前半期，伴随着他转向对统一场论的追求。在20世纪20年代初，爱因斯坦仍然把空间看成是一种独立的实在(1920a)，空间的几何结构可能具有一些场的性质(1922)。但是5年以后，爱因斯坦已经把度规关系看成是与场的性质同一的，并声称"广义相对论形成了场论纲领发展的最后一步"(1927)。

在这个转变过程中，爱因斯坦的"实用几何"观扮演着重要的角色。[102] 如果时空的几何结构只在测量杆与时钟的行为中显示自身，并且完全被以引力场为介质的引力相互作用所决定，那么就可以非常自然地把与物质以及引力场本身相互作用的引力场看作一种物理实体，而把几何结构看作它的结构性质，假如人们接受活动性、实在性与自我保存是实体的判据的话。值得注意的是，两篇已经澄清了爱因斯坦的几何观的特别重要的论文(1921a，1925b)，正好是在这段变化时期写成的。

因此，我们达到了最重要的结论。在进入思想发展的第二个阶段以后，爱因斯坦所做的不是将引力理论几何化，而是将时空几何引力化。[10]即他把几何看作引力相互作用的表现，而不是反过来把引力看作时空几何的表现。时空的几何结构被引力化以后，随着引力相互作用的演化而演化，其演化定律等同于场的动力学定律，即爱因斯

坦场方程。

涉及上述讨论的只是爱因斯坦关于引力场与具有几何结构的时空之间关系的观点。在爱因斯坦场论的前两个阶段，除了场的实在性以外，物质的实在性也是预设的。但是随着场论的发展，尤其是对真空场方程的深入考察，爱因斯坦被迫表示要考虑场与物质的关系。

必须解决的是如下的两难困境。一方面，爱因斯坦对马赫原理的承诺迫使他把物质看作是场的唯一来源，是场变化的唯一原因。另一方面，真空场方程数学上的有效解的存在似乎暗示着场本身是一种实体，而不是源自物质。如果他拒绝马赫原理，并把场看作一种基本物质，则他必须解释物质为什么像场方程显示的那样发生作用，成为场源以及场变化的原因。爱因斯坦的解非常简单。他把被他错误地还原为电磁场的物质，与引力场一起看作是同一基质即所谓非对称整体场的两种表现形式。这在概念上提供了两者相互转换的实物基础。因此他指出，“我确信，无论如何，几何场与其他种类场的差别不是逻辑地建立起来的”(1948b)。[11]通过这个解，爱因斯坦把几何纲领推向了一个新阶段，即统一场论的阶段。

[103] 总之，主要建立在广义相对论与爱因斯坦其他著作基础之上的几何纲领，在它所有三个阶段都与牛顿物理学截然不同，因为它是一种场论。在牛顿纲领中，引力表现为一种瞬时超距作用，并且不能解释为一种由单个物体决定并从中流溢出来的活动能力，而必须被看作只是通过虚空相互作用的两个物体之间的纽带。在爱因斯坦纲领中，引力被看作是一种局域的介质作用，不是瞬时地而是以光速传播的，被分开为一个物体的作用(单独被这个物体决定的场的激发)与另一个物体的反作用(由场引起的动量的时间变化)。在两个物体之间，场从一个物体到另一个物体传递动量与能量(Einstein，1929a，b)。

按照爱因斯坦的意见(1949)，马赫对牛顿理论的批评注定要失败，因为马赫预设了质量及其相互作用——而不是场——是基本的概念。但是，

> 战胜绝对空间概念或惯性系概念之所以成为可能，只是因为物质客体的概念渐渐被场的概念所取代，而成为物理学的基本概念……迄今为止，没有人发现任何可避开惯性系的方法，除非借助于场论的方式。
>
> (1953)

爱因斯坦是对的，他对牛顿理论的胜利是场论的胜利。但不幸的是，尽管他想以令人信服的方式证明物质与场(既有电磁场，又有引力场)能够用统一场论描述，但他失败了。

当然，爱因斯坦的场论是一种特殊的场论，其中场不可分割地与时空的几何结构联系在一起。那么，它与更富有成果的场论，即量子场论的关系是什么样的呢？这个有趣的问题，与大量历史与哲学的问题一起，将在本书后文提到。

5.3 进一步的发展

大约在爱因斯坦提交广义相对论的最后版本的同时，希尔伯特(David Hilbert)提出了一个物理学基本方程的新体系(1915，1917)，把米(Gustav Mie)的物质的电磁场论(1912，1913)与爱因斯坦的广义相对性综合起来。在他的新理论中，希尔伯特利用变分原理，从米的世界函数 $H(g_{\mu\nu}, g_{\mu\nu,\rho}, g_{\mu\nu,\rho\sigma}, A_{\mu}, A_{\mu,\rho})$ 出发，得到14个势的 [104] 14个方程：这些方程中的10个包含着与引力势 $g_{\mu\nu}$ 有关的变分，因

而被称为引力方程；而另外4个出自于与电磁势 A_μ 有关的变分，从而给出了广义麦克斯韦方程。希尔伯特声称，借助于在应用变分原理的过程中的数学定理，将在 n 个场中求得这4个关系：

> 4个[电磁方程]可以看作引力方程的结果……在这个意义上，电磁现象是引力效应。
>
> (1915)[12]

希尔伯特的引力与电磁的统一理论为几何纲领的早期发展提供了强有力的刺激。尤其是，他所设想的度规 $g_{\mu\nu}$ 与电磁势 A_μ 之间的联系，引导他通向物理学与几何学关系的一般观点：

> 物理学是四维赝几何学，其度规 $g_{\mu\nu}$ 与电磁量即物质有关。
>
> (Hilbert, 1917)

但是，希尔伯特既没有论及物理学的几何基础，也没有论及时空的几何结构。此外，电磁势 $A_{\mu\nu}$ 的几何对应物在他的理论中也是不清楚的。

对于以数学为基础结构的几何纲领的发展至关重要的是列维-奇维塔(1917)引入的矢量的无穷小平移概念。从这样一种概念出发，就能得到诸如黎曼曲率张量之类的黎曼几何中的基元。通过这种方式，黎曼几何被推广了。列维-奇维塔也赋予克里斯托费尔符号 $\Gamma^\mu_{\rho\nu}$ 三指标以意义，即表达仿射联络流形中线元的无穷小平移，其中线元的长度是不变的。因此，矢量在弯曲空间中的平行移动被设计成等价于协变微分的概念：$a^\mu_{;\nu} = a^\mu_{,\nu} + (\Gamma^\mu_{\rho\nu})a^\rho$。追随列维-奇维塔的思想，

黑森贝格(Hessenberg，1917)认为空间由大量通过平行移动紧密联系起来的小基元形成，即成为一个仿射联络空间。

爱因斯坦理论在完全几何意义上的第一个延伸是由外尔(1918a，b)给出的。受爱因斯坦早期思想的强烈影响，外尔坚持几何应被“视为一种物理实在，因为它作为作用于物质的实在效应的源头表现自己”，而且“引力现象也必须被放置在几何的说明中”(1922)。[105]而且，他想要提出一个理论，其中引力与电磁都来自于同一个源，不能任意地被分离，并且其中“所有物理量在世界几何中都有意义”(1922；参阅 1918a，b)。

为了这个目的，外尔必须推广爱因斯坦理论的几何基础。作为一个出发点，他把场论思想贯彻到几何学当中，而且作出了下面的批评：

> 黎曼几何只是走在达到纯无限小几何的理想的半途上。它仍然坚持根除“超距”几何的最后元素，即其往昔的欧几里得残余。黎曼假设在空间的不同点比较两个线元的长度也是可能的；不允许在“无限临近”几何中使用超距比较。单独一个原理是容许的；通过这，长度的区分就可以从一个点转移到无限临近它的另一个点。
>
> (1918a)

因此，在外尔几何中，长度的特定标准应当只用于它所在的时刻与位置，人们必须在空间与时间的每一个点设置单独的长度单位。这样一种单位标准系统称为规范系统(gauge system)。同样地，矢量或张量的概念只有在一个点上才是先验有意义的，而且只有在同一点

上才是可比较的。那么在全部时空中，它们的意义是什么呢？在这一点上，外尔吸收了列维-奇维塔关于无限小平行位移的思想以及黑森贝格关于仿射联络空间的思想。他自己原来的思想是这样一种假设：规范系统就像坐标系一样是任意的。

> $g_{\mu\nu}$只有在其等比例性的限度内才被 p 点的度规性质决定。在物理意义上，也只有 $g_{\mu\nu}$的比例具有直接的确实意义。
>
> （Weyl，1922）

外尔要求一个正确的理论必须具有双重不变性的性质：关于坐标的任何连续变换的不变性，以及在任何一种规范变换下的不变性。在规范变换中，$\lambda g_{\mu\nu}$被 $g_{\mu\nu}$取代，其中 λ 是位置的任意连续函数。外尔声称："这第二种不变性的性质的随附性(supervention)是我们理论的特点。"（同上）

规范系统的任意性要求矢量在不同点的长度 l 经历无限小变化，长度 l 能够表示为 $dl^2 = l^2 d\phi$，其中 $d\phi$ 是线性微分形式：$d\phi = \phi_\mu dx^\mu$（黎曼几何是当 $\phi_\mu = 0$ 时的极限情形）。这使得我们从二次 [106] 函数形式(黎曼线元) $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ 出发，就能得到

$$g_{\mu\nu,\rho} - \Gamma_{\mu,\nu\rho} - \Gamma_{\nu,\mu\rho} = g_{\mu\nu}\phi_\rho \qquad (1)$$

在 $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ 中用 $\lambda g_{\mu\nu}$取代 $g_{\mu\nu}$，人们就能轻易地得到 $d\phi'_\mu = d\phi'_\mu + d(\log\lambda)$。于是规范不变性暗示着，$g_{\mu\nu} dx^\mu dx^\nu$ 和$\phi_\mu dx^\mu$是与 $\lambda g_{\mu\nu} dx^\mu dx^\nu$ 及 $\phi_\mu dx^\mu + d(\log\lambda)$有共同基础的。因此，在反对称张量 $F_{\mu\nu} = \phi_{\mu,\nu} - \phi_{\nu,\mu}$ 中具有不变的意义。这个事实引导外尔设想在世界几何中，把 ϕ_μ解释为四势的，把张量 $F_{\mu\nu}$解释为电磁场。

从以上考虑出发，外尔坚持认为，不仅电磁场能够从世界几何中推导出来(1918a,b)[13]，而且时空的仿射联络 Γ 也依赖于电磁势 ϕ_μ 以及引力势 $g_{\mu\nu}$，如同我们从方程(1)能很容易看到的那样(1922)。 ϕ_μ 在时空的仿射联络 Γ 中的出现，暗示着时空的几何必定偏离黎曼几何。因此，外尔的统一理论不同于希尔伯特理论之处，既在于它更丰富的几何结构，又在于它在电磁理论与时空几何之间更清楚、更紧密的联系。在把世界几何作为本体论上基本的物理实在、而引力与电磁作为导出现象方面，外尔的理论也不同于爱因斯坦的理论，即它属于几何纲领的强版本。

外尔的长度移动的不可积性(non-integrability)概念，来源于长度单位的局域定义，并招来了很多批评。在批评者当中最著名的有爱因斯坦，他指出了它与观测到的谱线的确定频率相矛盾(Einstein, 1918b)。作为回应，外尔建议由"通过校准来确定"的概念取代"固定不变"的[长度]概念，来解释并消除这个困难(1921)。但这过于思辨，而且他自己在细节上也没有完成它。

虽然这个概念面临着所有上述困难，但另一个与之联系(或者来源于它)的概念，即(可扩容的)规范不变性的概念，却具有重要的启发意义。在他的统一理论中，外尔建立了电荷守恒与可扩容规范不变性之间的联系。他把这个联系看作"支持我的理论的最强有力的一般论证之一"(1922)。虽然可扩容规范不变性的原始思想在它出现不久就被放弃了，因为它的推断在新的量子力学语境中与观测相冲突，但外尔还是在1929年复兴了这个思想(见9.1节)。这次，局域不变性是电磁学中量子相位的不变性。我们目前关于电荷与电磁场的观点，在很大程度上依赖于这个思想。实际上，物理定律的规范 [107] 不变性与时空的几何结构之间的关系，已经是当代物理学中最令人着

迷的主题之一(见第三篇)。

但是，在爱丁顿看来，外尔的几何仍然遭受了不必要的限制。爱丁顿希望证明：

> 一旦把外尔的几何从这些限制中解放出来，整个框架变得简化了，而新的光芒将照耀着物理学基本定律的起源。
>
> (1921)

在借助广义相对论来超越欧几里得几何的过程中，爱因斯坦得到了引力；在借助规范不变性原理来超越黎曼几何的过程中，外尔得到了电磁学。因此，人们会询问，通过进一步的推广还能得到什么。爱丁顿的回答是，我们希望能够得到非麦克斯韦的结合力，它能够与库仑排斥力相抗衡，从而把电子聚合在一起。

像外尔一样，爱丁顿从平行位移与规范不变性的概念出发。他们之间的主要差异存在于以下事实中：外尔采用 $dl^2 = l^2 d\phi$ 与方程(1)，而爱丁顿则把方程(1)的右边写作 $2K_{\mu\nu,\rho}$，而非特殊形式 $g_{\mu\nu}\phi_\rho$。这导致 $dl^2 = 2K_{\mu\nu,\rho}\xi^\mu\eta^\nu dx^\rho$。于是，按照以下表示：

$$S_{\mu\nu,\sigma} = K_{\mu\nu,\sigma} - K_{\mu\sigma,\nu} - K_{\nu\sigma,\mu}\text{，以及 } 2k_\mu = S^\sigma_{\sigma\mu}\text{，}$$

广义规范不变曲率张量 $G^*_{\mu\nu}$ 被分成了对称部分与反对称部分：

$$G^*_{\mu\nu} = R_{\mu\nu} + F_{\mu\nu}\text{，其中 } F_{\mu\nu} = k_{\mu,\nu} - k_{\nu,\mu}\text{，}$$

当 $F_{\mu\nu}$ 能够被看作电磁场时，对称部分

$$R_{\mu\nu} = G_{\mu\nu} + H_{\mu\nu}\text{，}$$

其中 $H_{\mu\nu} = k_{\mu,\nu} + k_{\nu,\mu} - (S^\sigma_{\mu\nu})_\sigma + (S^\beta_{\alpha\nu}S^\alpha_{\beta\mu} - 2k_\alpha S^\sigma_{\mu\nu})$ 包括描述引力

的曲率 $G_{\mu\nu}$，以及表示整个能量张量与电磁能张量之间差别的 $H_{\mu\nu}$。爱丁顿设想这个差别必定代表整个能量张量的非麦克斯韦电学部分。

虽然爱丁顿得到了更一般的几何学，但他坚持实在世界的自然几何是黎曼几何，而不是外尔的广义几何，也不是他自己的几何。他指出：

> 我们所寻找到的不是实际空间与时间的几何，而是世界结构的几何，这个世界结构是空间、时间与事物的共同基础。
>
> （同上）

毫无疑问，引入世界结构的几何概念对于几何纲领的发展，以及几何学概念的演化都是重要的(见 11.3 节)。 [108]

爱因斯坦在精神上与外尔和爱丁顿一脉相通并部分地受到两人论文的激励，提出了建立在反对称联络或“度规”基础上的统一场论(UFT)。[14]除了得到标准广义相对论的赝黎曼时空以外，他也致力于描述电磁学的一些更深入的几何结构。虽然在数学上相当复杂，但爱因斯坦的统一场论在物理思想上是相对简单的：除了代表引力的时空的度规结构外，必定也存在代表电磁力的一些其他时空结构。但是，根据爱因斯坦的意见，存在着两种彼此独立的时空结构的思想是不能容忍的。所以，人们应当寻求在其中这两种结构组成一个统一整体的时空理论(Einstein，1923a，1930a)。从这个基本思想出发，我们能够看到爱因斯坦的统一场论如同他的广义相对论一样，显示出鲜明的几何印象，实际上是几何纲领的另一个阶段。

在爱因斯坦(1945 与 1948b)关于统一场论的著作中，整个统一场由一个复厄米张量 g_{ik} 来描述。这非常不同于广义相对论中由对称张

量 $g_{\mu\nu}$ 来描述的引力场。 统一场(或势)可以被分为对称部分与反对称部分:

$$g_{ik} = s_{ik} + \mathrm{i}a_{ik}, \text{ 其中 } s_{ik} = s_{ki}, a_{ik} = -a_{ki}$$

这里 s_{ik} 是能够等同于度规的对称张量或引力势，而 a_{ik} 是能够等同于电磁场的反对称张量。 作为三十年寻求统一场论的最后结果，爱因斯坦通过放弃无限小位移场 $\Gamma^{\rho}_{\mu\nu}$ 必定对其下标是对称的这一限制，推广了他的引力理论。 采用这种方式，除了一个对称部分(引向纯引力场理论)，爱因斯坦还得到了 $\Gamma^{\rho}_{\mu\nu}$ 的反对称部分(可望引向电磁场理论)(1954b)。

因此，这个理论在美学上是令人满意的，引力场与电磁场构成了同一个统一场的两个部分。 但遗憾的是，爱因斯坦的统一场论得不到经验事实的支持，显然也没有类似于广义相对论从惯性质量与引力质量的等价性那里得到的那种支持。 迄今为止，缺少支持的理由远比没有合适的实验更为深刻。 在数学上，正如爱因斯坦自己承认的那样，“我们完全没有任何方法得出系统的解……出于这个理由，我们现在不能把非线性场论的内容与经验进行比较”(1954b)。 从我们 [109] 目前关于基本相互作用的统一的知识来判断，爱因斯坦统一场论的基本物理思想，诸如非引力场与时空的几何结构之间的关系，不同部分场之间的结构关系，部分场与总体场之间的结构关系，都是没有经验基础的纯思辨的东西。

在 1923 年到 1949 年之间，爱因斯坦花费了很大精力，但在统一场论方面没有作出主要的突破。 而在同一个时期，电磁学变成了第一个完全狭义相对论性量子场论，没有为人所知的直接几何特征。后来的发展也证明了，弱相互作用与至少某些特定的强相互作用，看

来也需要采用这种没有直接几何特征的相对论性量子场论来描述。[15]自从 20 世纪 70 年代中叶以来，量子场论的成功已经刺激了统一场论思想的复兴(见 11.2 节)。但是，在复兴了的统一框架中，总体场与部分场之间的结构关系，以及不同部分场之间的结构关系，已被证明比对称部分与反对称部分的组合复杂得多。大量的新思想，诸如分级的对称性破缺与希格斯机制(Higgs mechanism)，对于理解这些关系是必不可少的。因而在内在严密性与连贯性方面，爱因斯坦的统一场论几乎不能被视为成熟的理论。

值得注意的是，根据爱因斯坦的时空观，在他的统一场论中，从无限小位移场构造出来的几何结构描述了总体场的结构性质，而不是作为潜在的物理实在的时空结构。在本体论上，后者已经被毫无保留地还原为场的结构关系。所以，如同惠勒(1962)声称的那样，在爱因斯坦的统一场论语境中，"统一"所意味的不过是两种场的统一，而不是物理场与时空的统一。

一种试图改变标准广义相对论的黡黎曼结构的略为保守的方法是由嘉当(Elie Cartan，1922)提出的。根据嘉当，描述具有内在角动量的物质分布的时空模型，应当由具有与自旋密度相联系的挠率的弯曲流形来表示。嘉当的基本思想被一些作者发展了。[16]可以设想，挠率可以在那些类似中子星的对象内部产生可观测的效应，中子星本来就有强磁场，并可能伴随着自旋密度的一个实平均值。

与引入了变分 dl^2 并把它和电磁势的存在联系起来的外尔相比，卡卢察(Theodor Kaluza，1921)停留在度规几何的王国里，通过扩张宇宙的维度来寻求将电磁场包括在内[的统一场论]。卡卢察的 [110] 五维几何的线元能够被写成 $d\sigma^2 = \Gamma_{ik}dx^i dx^k (i, k = 0, \cdots, 4)$，这里 Γ_{ik} 是五维对称张量的 15 个分量，它们与四维 $g_{\mu\nu}(=\Gamma_{\mu\nu}, \mu, \nu =$

1，…，4）与电磁势 A_ν（$=\Gamma_{0\nu}$，$\nu=1, \cdots, 4$）有关。 那么第 15 个量 Γ_{00}意味着什么呢？ 根据奥斯卡·克莱因（Oskar Klein，1927），我们能够尝试把 Γ_{00}与表示物质的波动函数联系起来，以便物质与场的形式统一能够实现。 在这个意义上，Γ_{ik}能够被看成是爱丁顿的世界结构几何学的一种实现。

为了把量子效应整合到五维理论中，奥斯卡·克莱因追随德布罗意的思想（见 6.5 节），猜想作用量子可能从第五维的运动的周期性中产生。 这些运动在普通实验中不是可察觉的，因此我们可以对整个运动平均。 奥斯卡·克莱因（1926）甚至声称，"作为场方程的几何基础的激烈修正"，把第五维引入到物理图景中"是由量子理论提出的"。

虽然奥斯卡·克莱因本人（1927）很快认识到他在把量子现象整合到时空描述方面失败了，但维布伦（Oswald Veblen）与霍夫曼（1930）并不因此泄气。 他们建议以四维的射影几何（projective geometry）取代基于卡卢察—克莱因五维理论的仿射几何，而且证明了当施加于基本射影张量的限制（它被迫将射影理论还原为奥斯卡·克莱因的仿射理论）被放弃时，将会得到包含克莱因—戈尔登型波动方程的一组新的场方程。 在这个数学证明的基础上，他们过分乐观地声称，相对论中射影几何的使用"似乎使得把波动力学带入相对论框架中成为可能。"

甚至在目前，卡卢察—克莱因理论的基本思想也没有消亡。 已经作出了很多尝试以用来证明，表征内部（动力学）对称性的规范结构是高维时空的几何结构的表现（在 11.3 节中有更多的讨论）。

惠勒是几何纲领最积极的鼓吹者之一。 他在 1962 年发表了有影响的著作《几何动力学》（*Geometrodynamics*），在该著作中他把几何

看作原初实体，认为引力不过是几何的一种表现，而其他所有一切均来源于几何，或者可从几何构造出来。他相信在时空几何本身内部能够存在负载能量的波，时空几何预设了几何实体是一种物理实在。[111] 也就是说，几何实体通过它而实在化了（Misner，Thorne，and Wheeler，1973）。

关于量子现象，惠勒的立场是相当适度的。他没有从时空几何中推导出量子理论的野心。相反，他把量子原理看作在物理学的建构中比几何动力学更基本的东西。他把量子原理补充到几何动力学中，并因而重新表述了几何动力学。这种立场导致了一些意义深远的结果。一开始，它把惠勒引向量子几何动力学。但在最后，它讽刺性地把惠勒引向否定几何纲领。

根据量子几何动力学，在几何小距离中存在着量子涨落，这导致了多连通空间(multiply connected space)的概念。出于这些理由，惠勒论证道，空间必然具有泡沫状结构。他把电看作是陷落到多连通空间中的力线，而把电荷在自然界中的存在视为空间在小距离中是多连通的证据。根据他的观点，一个粒子能够被看作一种“几何动力学激发，而各种场同样能够根据多连通几何的激发模式来解释”(1964b)。

但是，“量子涨落”作为他宇宙的几何图景的基本元素，也悖论式地逐渐毁损了这幅图景。量子涨落使得连通性的变化成为必要。这是与微分几何的思想不能并存的，微分几何预设了点邻域(point neighborhood)的概念。随着微分几何的失效，宇宙的几何图景也失败了：它只能提供在最小距离发生的事态的粗糙近似。如果几何不是物理学的终极基础，那么必定存在一种实体——惠勒称之为“前几何”(pregeometry)——它既比几何又比粒子更原始，并且作为两者建

立于其上的基础。

因而问题在于，几何与物质都成为其表现的前几何究竟是什么。惠勒的回答是：一种原始的与根本的混沌(chaos)(1973)。实际上，惠勒对建立在量子原理基础上的前几何的理解，超越了几何纲领的视野。惠勒思想的演化，展示了在几何纲领中整合量子原理的内在困难之一，一种调和离散与连续的困难。

彭罗斯是几何纲领的另一个积极鼓吹者。他思想的演化是特别有趣的。起先，彭罗斯(1967a)提出了时空结构的旋量方法(spinor approach)。最简单的基砖(从它出发，标准场论中所有张量场与旋量场的值都能构造)是二分量旋量；根据这一事实，彭罗斯提出，我[112]们能够在一开始就把时空看作这类旋量场的载体，从旋量场的作用中推导出时空的结构。因此，他既得到了时空的赝黎曼结构，也得到了时空的旋量结构，后者的存在是必需的，因为自然界中存在费米子。

随后，彭罗斯(1967b，1975)提出了更雄心勃勃的理论——扭量理论(twistor theory)——来沟通时空几何学与量子理论。他注意到存在着两种连续统：四维的实连续统代表着时空的舞台，世界的现象被设想为在其中发生；而产生概率幅概念与叠加定律的量子力学的复连续统，引向了描述量子现象的复希尔伯特空间的图景。通常，量子力学的思想是被简单地叠加到四维时空的经典图景中。但是，扭量理论的直接目的是提供一个物理学框架，其中这两种连续统合而为一。

彭罗斯发现他的目的是可以达到的，因为在旋量群 SU(2，2)与闵可夫斯基时空 $C_{+}^{\uparrow}$（1，3）的 15 参数共形群之间存在着局域同构，这种同构使得由定义了扭量空间的结构的两个旋量部分构成的复数，

必须与时空的几何紧密联系在一起，并且也作为量子力学的概率幅以不同的姿态出现。

但是，在彭罗斯的理论中，表征处在自由运动中的经典无质量粒子的扭量是基本的，而一开始在理论中没有出现的时空点被看作是导出的对象。这暗示着，根据彭罗斯，“连续统概念最终有可能从物理理论的基础上根本消除”（1975）。但是，消除时空连续统损害了几何纲领强版本的本体论基础。这里，我们发现了几何纲领强版本的另一个叛逆，虽然他的理论动机与智力探险完全不同于惠勒。

5.4 拓扑学研究：奇点、视界与黑洞

令人好奇的是，彭罗斯在本体论层次上放弃几何纲领强版本，是与在描述性层次上放弃它紧密相连的，直接受其激励，甚至是不可抗拒地受其支配，通过他把几何纲领推向新的阶段，也就把它推向了尽头。更具体地说，彭罗斯与霍金及其他物理学家一起，通过他们对强引力场的行为以及这些场作用于光与物质的效应的拓扑学研究[17]，[113] 已经成功地在广义相对论的表述中澄清了奇点的本质，显示了时空的微妙结构，并因而帮助发展了20世纪60年代末以来广义相对论的一种新解释。但是，这些成就也显示了广义相对论的内在局限。因此，不得不求助于量子原理，以便能够保持发展世界的协调图景的希望。

有三个概念对于广义相对论的拓扑解释是关键的：奇点(singularity)、视界(horizon)与黑洞(black hole)。

奇点(A)

从广义相对论被表述的时候开始，物理学家就长期为出现在场方

程的某些解中的奇点所困扰。如果我们看一下施瓦氏解

$$ds^2 = (1 - 2Gm/rc^2)c^2dt^2 - dr^2/(1 - 2Gm/rc^2) - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (1)$$

（这里 m 是点质量或远离其他物体的无自旋的球对称星体的质量，而 G 与 c 是牛顿引力常量与光速），我们很容易发现，当 $r = 2Gm/c^2$ 或 $r = 0$ 时，度规不能很好地定义。当 $r = (3/\lambda)^{1/2}$ 时（或者在一种转换过的表述中，当 $r = \pi R/2$ 时；见 4.4 节），这同样发生在德西特解中

$$ds^2 = (1 - \lambda r^2/3)dt^2 - (1 - \lambda r^2/3)^{-1}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (2)$$

当涉及这些度规奇点的本质与解释时，物理学家意见有分歧。爱因斯坦(1918a)认为，德西特度规中的奇点显示了德西特世界是“一个物质完全聚集在 $r = \pi R/2$ 表面上的世界”。外尔(1918a)把德西特奇点解释为一种不可达的视界，但坚持“至少在该视界上必定存在着质量”。他与爱因斯坦持有相同的观点，即世界只有达到奇点才能够是空质量的。

爱丁顿对于度规奇点的解释非常不同于爱因斯坦与外尔的解释。对于他来说，度规奇点“未必象征物质粒子”。这方面的理由是，“我们能够通过变换坐标来引入或消除奇点。无法知道应当指责的究竟是世界结构还是坐标系的不适宜性”(Eddington，1923)。

爱丁顿关于度规奇点只是坐标奇点的观点被他的学生勒梅特详尽 [114] 描述。勒梅特(1932)是给出施瓦氏奇点的坐标本质的严格数学证明的第一人。在这个证明的基础上，他得出结论：“施瓦氏场的奇点因而是虚构的奇点，类似于德西特宇宙的原始形式中的中心视界的

奇点。"

爱丁顿对度规奇点的另一种解释也很有影响：奇点构成了一个不可穿透的球，一个魔环，物质与光聚集其上，但不能穿透。他把德西特奇点（处于空间中 $r = \pi R/2$ 处的有限距离中的非连续性）看作物理上不可达的视界，而且指出在那里，

> 光，就像其他一切东西一样，被静止地约束在时间停顿的区域，而且它永远不能逃离这个世界。超越这个距离（$\pi R/2$）的区域总是通过这个时间之垒对我们关闭着。
>
> （Eddington，1918）

在他富有影响的著作《相对论的数学理论》（*The Mathematical Theory of Relativity*，1923）中，爱丁顿进一步发展了他的度规奇点的视界观。通过对广义解：

$$ds^2 = (1 - 2Gm/rc^2 - \lambda r^2/3)c^2 dt^2 - dr^2/(1 - 2Gm/rc^2 - \lambda r^2/3) - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (3)$$

（其中施瓦氏解或德西特解能够通过令 $\lambda = 0$ 或 $m = 0$ 而得到）作出评论，爱丁顿指出：

> 在 g_{44} 消失的地方[即奇点出现之处]存在着一个不可穿越的垒，因为 dr 的任何变化都对应于借助测量杆测量得到的一个无限大距离 ds。g_{44} 的两个正立方根近似地为 $r = 2Gm/c^2$ 与 $r = (3/\lambda)^{1/2}$。第一个根[施瓦氏奇点]将代表粒子的边界，并……给予它不可穿透性的表象。第二个根[德西特奇点]是在一个非常大的

距离上，并可以被描述为世界的视界。

(1923)

视界

在德西特模型中，$r = \pi R/2$ 处的表面的视界本质很快就被接受了。这个表面被视为是在有限的但物理上不可达的空间距离上的时间表面，因为在那里所有物理现象已经不再具有时间延续，也没有任何东西能够到达。在施瓦氏解中，$r = 2Gm/c^2$ 是在有限时间内不可达与不可穿透的，这个观点也被人们接受了——因为当 r 接近 $2Gm/c^2$ 时，光的波前倾向于保持静止（或等价于无限大红移），而所有事件倾向于被推向无限——但这种观点不是没有挑战者。

[115] 挑战这种观点的概念基础是勒梅特关于施瓦氏模型的非奇点解（1932）。只有通过这种对施瓦氏奇点的虚假本质的理解，关于通向视界形成的物理过程的有意义的研究才成为可能，对视界的拓扑性质的数学澄清才可验证。似乎不可穿透的视界的存在的物理理由，是由钱德拉塞卡（Subrahmanyan Chandrasekhar，1935），茨维基（Fritz Zwicky，1935，1939），奥本海默与沃尔科夫（Robert Oppenheimer and George Volkolf，1939），以及奥本海默与斯奈德（Oppenheimer and Hartland Snyder，1939）等，在研究星体演化、引力坍缩（爆聚）与到达临界圆周（即黑洞的形成，见下面）的过程中，于20世纪30年代提出的。但是，这些研究的大部分只挑战施瓦氏视界的不可达性，而不是不可穿透性。正是罗伯逊（Howard Robertson，1939）在1939年论证道，勒梅特的非奇点解允许对穿越视界（$r = 2Gm/c^2$）并直达施瓦氏球心（$r = 0$）的任何粒子或光子的轨迹进行完美的规则描述，因而否定了不可穿越的视界这个教条。

罗伯逊注意到这种描述令人困惑的含义：观测者将永远不能看到粒子达到 $r = 2Gm/c^2$（当粒子接近 $r = 2Gm/c^2$ 时，观测者接收到越来越多的来自粒子的红移光；即在 $t = \infty$ 时，粒子穿越视界），虽然粒子是在有限的固有时经过 $r = 2Gm/c^2$ 并到达 $r = 0$。因此，似乎有两个不相容的观点：从外部观测者的立场来看，在视界处时间停顿，事件冻结。这似乎暗示着视界不能被看作物理上真实的，因而广义相对论在视界上破缺了。但是，一个下落粒子将会穿越视界不断下落，既注意不到任何时钟的变慢，也看不到视界上的无限红移或任何其他反常效应。这意味着从下落粒子的立场来看，视界不会对广义相对论提出任何挑战。正是芬克尔斯坦（David Finkelstein, 1958）引入了新的参考系，使两种观点协调起来了。

对于物理学家，视界的意义一直是非常不清楚的，直到 1956 年林德勒（Wolfgang Rindler）将粒子视界（PH）定义为在任何宇宙瞬间把所有粒子（包括光子）区分为两类的光波波前：那些已处在我们视野中的粒子为一类，而所有其他粒子为另一类。根据林德勒，粒子视界不同于事件视界（EH），在事件视界中，一个球形光波波前向我们会聚而来，把在每一条测地线上所有穿过我们的实光子与虚光子分为两类：那些在有限时间内到达我们的光子，以及在有限时间内不能到达我们的光子（Rindler, 1956）。林德勒的定义是精彩的。但是，视界的含义远比任何已经详细研究的定义更复杂。人们不可能在不提及黑洞概念的情况下澄清视界的含义。 [116]

黑洞

虽然黑洞一词是晚至 1967 年年底才首先由惠勒提出来的，但是黑洞的思想绝不是新鲜的东西。早在 1783 年，米切尔（John Michell）从牛顿光的微粒说出发，在皇家学会宣读的一篇论文中，就

提出了可能存在临界圆周(其中没有任何光能够逃离)与“暗星”(dark bodies，正如20多年后拉普拉斯称呼它们的那样)的思想，文章设想：

> 光粒子像所有其他物体一样以同样的方式被吸引……所以，如果与太阳同样密度的星球的半径以500比1的比例超过太阳的时候，所有从这样的天体发射出来的光，都将由于本身的固有引力而被迫返回天体。
>
> (Michell，1784)

在19世纪初光的波动理论取代光的微粒说以后，米切尔的思想被遗忘了。人们简直无法想象引力是如何作用于光波的。

在广义相对论语境中，米切尔的思想通过施瓦氏解复活了，施瓦氏解代表着在真空中单质量中心(即一个星体)的引力场，并决定着环绕该中心的几何学行为。这个解预言，对于每一个星体存在着一个临界圆周，其值只依赖于它的质量。可以证明，施瓦氏解中的临界距离 $r = 2Gm/c^2$，即所谓的施瓦氏奇点或施瓦氏半径，不过是米切尔的星体的临界圆周，在这个距离以下没有光能够逃脱其表面。或者，用现代术语说，它是一个事件视界。

虽然关于大质量星体的临界圆周的数值估计在这两种情形中是相似的，但数值估计所依据的概念非常不同。(i)根据米切尔的牛顿概念，空间与时间是绝对的，光速是相对的；而施瓦氏解预设了相反的立场；(ii)按照米切尔的观点，光粒子能够飞离临界圆周一点点，当光粒子升高时，其速度将会由于星体的引力而降低，最后它们将被拉回星体。因此，对于一个靠近星体的观测者，通过慢行的光看到星

体是可能的。但是，在施瓦氏解的情形中，从事件视界发出的光必定是无限红移的，因为那里时间的流逝无限膨胀了。因此，对于一个视界外的观测者，来自星体的光是不存在的。光存在于星体内部，但它不能逃离视界，而且必定向中心运动。(iii)从牛顿的立场来看，小于临界圆周的星体的稳定性能够维持，因为引力挤压被其内部压力所制衡。但是，根据现代观点，当星体的核能源耗尽时，没有内部压力能够抗衡引力挤压，星体将在自身的重力作用下坍缩。 [117]

第一个注意到施瓦氏解的上述含义的人是爱丁顿(1923)。在他的著作《恒星的内部结构》(*The Internal Constitution of the Stars*, 1926)中，爱丁顿进一步概括出在临界圆周以内会发生的事情：

> 首先，引力是如此巨大，以至于光不能从星体中逃离，光线就像石头落向地球一样落向星体。其次，光谱线的红移是如此巨大，以至于光谱被移到了不存在的地步。第三，质量产生的时空度规曲率如此巨大，以至于环绕星体的空间将关闭起来，而把我们留在星体外面(即没有办法知道里面的情况)。
>
> (1926)

他的结论只是"类似那样的事情不可能出现"。

爱丁顿的观点受到他的印度学生钱德拉塞卡(1931, 1934)的挑战，后者在量子力学所隐含的简并态的存在的基础上，论证了一颗大质量恒星，一旦耗尽了其核能源，就会坍缩；而且这种坍缩必定无限制地进行下去，直到引力变得如此强，以至于光不可能从星体中逃离。他也论证道，这种爆聚暗示着星体的半径必定趋向于零。爱丁顿(1935)拒绝了他的学生的上述论证："这几乎是相对论简并公式的

归谬法……我认为应当存在着一种自然律把星体从这样一种荒谬的形式中解救出来。”

但是，爱丁顿的拒绝不具有逻辑说服力，而对引力爆聚含义的研究仍在继续。茨维基(1935，1939)证明了，当爆聚星具有较小质量时，将会引起超新星爆发，导致中子星的形成。当爆聚星质量比中子星(两倍太阳质量为其上限)质量大得多时，奥本海默与斯奈德(1939)论证道，这样形成的中子星将是不稳定的，爆聚仍将继续，星体的尺度将会缩到小于临界圆周(而形成黑洞)，直到它变成具有零体积与无限大密度的一个“点”。也就是说，爆聚将在黑洞中心终结于一个奇点(具有无限大的潮汐引力)，这个奇点会毁灭与吞没落入黑洞的一切东西。(大质量旋转星体的坍缩结果是具有外部度规的黑洞，它最后将变成由场方程的克尔旋转对称解来描述的黑洞。见 Kerr，1963。)

[118]

在 20 世纪 60 年代，黑洞被看作空间中的一个洞，东西可落到洞里面，却没有什么东西可从洞里面涌现出来。但是，自从 70 年代中期以来，人们开始认识到黑洞不是空间中的洞，而是一个除了自旋、质量与电荷之外没有任何其他物理性质的动力学对象。

黑洞，即 $r < 2Gm/c^2$ 的区域，是一个具有巨大密度与巨大引力场的区域。因此，黑洞的研究成了在巨大引力场背景中检验广义相对论的便捷方式。既然广义相对论的一致性只有当它能够完成其逻辑推论并且其最后结果得到检验的时候才能被澄清，那么黑洞的概念情景的澄清对于检验广义相对论的一致性是至关重要的。但是，没有对处于黑洞中心的奇点本质的清楚理解，澄清黑洞概念将是不可能的。

奇点(B)

奇点本质的合理理解不可能在新牛顿式的框架内得到，其中数学上黎曼时空流形的采用，是基于空间与时间的牛顿构想。在这个框

架内，奇点是在函数理论的意义上定义的。即当一个引力场在一个特定的时空点不可定义或变得奇异时，它被说成是具有一个奇点。但是，在广义相对论的框架内(见 4.2 节)，一个时空点本身必须被场方程的解定义，因为度规场是几何的原因，并由时空构成。如果度规在 P 点可以定义，那么点 P 只能属于时空结构。世界不能包含在其中度规场不可定义的点。因此，在广义相对论的框架内，声称度规在一个时空点是奇异的(类似于前广义相对论场或函数理论中的奇异点)，是没有意义的。

根据霍金与彭罗斯(1970)，时空奇点必须根据测地线的不完备性来定义，而一个无奇点的时空是测地线完备的时空。如果一条曲线不可延伸，那么它就是不完备的。一条不完备的类时测地线隐含着，在固有时的有限间隔之前(或之后)，能够从虚无中出现(或消失 [119] 于虚无中)的检验粒子是存在的。因而奇点是时空中缺少的某种东西。但是，这种缺少的某种东西与它跟其他时空区域的关系，仍然能够由陷入时空的理想点来表示，虽然这些理想点本身不是世界点或可能事件的位置(Schmidt, 1971)。

很多人论证道，压碎物质的奇点与时空几何是广义相对论天真地应用于强场(即大爆炸或引力坍缩)的结果；虽然大爆炸是被宇宙膨胀与宇宙微波背景辐射的观测所支持的，而坍缩是被中子星的观测所支持的。例如，爱因斯坦就拒绝了奇点的思想，并坚持：

> 对于高密度的场与物质，人们不可以假设[引力场]方程的有效性，而且人们不可以得出如下结论："膨胀的开端"必定意味着数学意义上的奇点。
>
> (1956)

彭罗斯与霍金作了不同的论证。借助于一些关于宇宙整体性质的显然合理的假设——诸如(i)场方程的有效性；(ii)负能密度的不存在；(iii)因果性；(iv)类似这样的 P 点的存在：所有通过 P 点指向过去的类时测地线在 P 点的过去的紧致区域再次开始会聚(即这样一个思想论断，宇宙中的物质所产生的引力充分相互吸引并足以产生奇点)——并利用拓扑学方法，他们成功地证明了无论是坍缩星体，还是宇宙的演化，都将不可避免地导致奇点。

引入拓扑学方法探索施瓦氏时空最隐秘的结构，并因此能够检查黑洞内部时空的第一人是克鲁斯卡尔(Martin Kruskal，1960)。但却是彭罗斯系统地引入整体概念，并与霍金一起，利用它们得到了关于奇点的结果，而且这个结论不依赖于宇宙的物质内容的任何精确对称性与细节，即著名的霍金—彭罗斯奇点定理(Hawking-Penrose singularity theorems)。

在上述假设当中，(ii)是非常弱的与合乎情理的条件，(iii)的打破将是我们物理推理的整体崩溃，而(iv)不仅被微波背景辐射的观测所支持(Hawking and Ellis，1968)，而且也被如下的合情推理所支持。如果星体的引力强到足以形成视界，把要往外跑的光拉回到里[120]面，那么在这一切发生以后，没有什么东西能够阻止引力变得越来越强，以至于星体将继续不可抗拒地爆聚为零体积与无限大密度，于是创造并融合为一个奇点，其中时空曲率或潮汐力变得无穷大，而时空将停止存在。这意味着任何黑洞本身内部必定都有一个奇点，因为它总是具有一个视界。因此，广义相对论所要求的坍缩星体中物理奇点的出现和宇宙的膨胀，正如奇点定理所证明的那样，已经对广义相对论本身的有效性提出了最严肃的概念问题。

在数学上，施瓦氏线元的两个奇点($r = 2Gm/c^2$ 与 $r = 0$)具有

不同的拓扑学意义。由 $r = 2Gm/c^2$ 代表的奇点(它在物理上能够被解释为一种事件视界而非真正的奇点),提出了在一个球面边界两侧的不连续性问题,而且暗示着我们关于世界的因果联系概念的剧烈修正。由 $r = 0$ 代表的另一个奇点,是一个真正的奇点,代表着时空的尖锐边缘,超越了这个边缘就不存在时空。既然霍金—彭罗斯奇点定理已经排除了关于大爆炸宇宙学的反弹模型,那么时间的经典概念在过去的奇点处必定有一个开端,而当星体已经坍缩时,至少部分时空将会终结。

在物理上,奇点处的引力如此巨大,以至于那里的所有物质都被毁灭,时空是如此弯曲,以至于时间本身将停止存在。奇点处的无限大引力是广义相对论定律必定在黑洞的中心(奇点)破缺的一个清楚信息,因为没有任何一种计算将是可行的。而这意味着广义相对论不是一个普适有效的理论。

因此,为了拯救广义相对论的一致性,必须发现一种能够中断大爆聚的机制。考虑到近奇点处极其高的密度与极其小的距离,量子效应必定在那里的物理过程中扮演重要的角色。因此,物理学家几乎一致的意见是,只有量子引力理论(广义相对论与量子力学的美满结合),才能提供必要的机制阻止奇点的出现,拯救广义相对论的连贯性,虽然这也意味着广义相对论将不再被视为完备的基本理论。

量子力学对于广义相对论问题的第一个成功应用,而且直到现在为止还可能是唯一的应用,是霍金对于黑洞热辐射的量子力学计算(1975)。这些年代里,有大量关于引力场量子涨落和概率的量子泡沫的华丽辞藻(见 Hawking, 1987)。但是,缺少一个一致的量子引 [121] 力理论,已经迫使物理学家诉之于彭罗斯的宇宙监督猜想(cosmic censorship conjecture),这个猜想宣称一个奇点总是被一个事件视界

所包围，因而对于外部世界来说总是隐蔽的(Penrose，1969)。虽然想要发现真正反例的所有尝试都失败了，但是这个猜想的任何证明还没有人给出。

注释

1. 关于这个主题的历史说明，见 Earman(1989)。

2. 绝对运动的存在未必与关系论不相容，因为与定义绝对运动有关的时空结构，仍然能够被证明在本体论上是由诸如广义相对论情形中的引力场之类的特定实体构成的。

3. 牛顿自己的观点复杂得多了。可以用引文来证明马赫的解释是能够被反驳的。例如，牛顿写道，“空间绝对不能在自身中存在”(1978，p.99)，以及“空间是事物对事物的作用……没有东西存在或能够存在，而不以某种方式涉及空间……空间是第一存在物的流出效应，因为如果任何一种东西都被设定为存在，那么空间也被设定为存在”(同上，p.163)。厄曼(John Earman)声称牛顿拒绝实体主义教条的任何形式，即“空间的存在不依赖于其他任何事物的存在”(1979)。但是，无论实际上牛顿采取什么样的立场，马赫对牛顿立场的解释与批评，在爱因斯坦关于物理学基础的观点成型的时候，是富有影响力的，而且被爱因斯坦全身心地接受。

4. 这里出自牛顿的引文也能用来反驳马赫对牛顿立场的解释。例如，牛顿在“论引力”(De gravitation)中宣称空间“并不承受那些主导实体的有特征的影响，即作用，诸如在心灵中的思想和在物体中的运动”(1978，p.99)，因此显然拒绝了空间是活动的或是活动的源泉的观点。

5. 值得注意的是，几何的这种实用观点与牛顿的意见一致，虽然关于空间的根本观点是与之完全不同的。牛顿指出，“几何是在力学实践中发现的，而且不过是普适力学的一部分”(1678，in 1934)。

6. 庞加莱写道，当天文观测与欧几里得几何学不一致的时候，“有两条道路向我们开放：我们可以要么重新审视欧几里得几何学，要么修正光学定律，并假定光线不是严格按照直线传播的”(1902)。

7. 格林鲍姆坚持这种关于爱因斯坦观点的立场，而这是被施泰因拒绝的。格林鲍姆与施泰因的激烈争论被厄曼、格利穆尔和施塔赫尔(Earman，Glymour，and Stachel，1977)所记载。这里，我只是试图概括我自己对于爱因斯坦关于(编时)几何学与测量仪器之间关系的观点的理解。

8. 见附录 A1。

9. 即使爱因斯坦的确不知道超新星作为引力辐射之源的可能性，但这确实暗示着引力场与其他场的相互变化。

10. 这个结论被爱因斯坦自己的论断支持：“我不同意广义相对论是引力场物理学几何化的思想。”(1948a)

11. 这里由于时空几何与引力场的紧密联系，几何场意味着引力场。

12. 对希尔伯特错误结论的批评，见 Stachel(1991)。希尔伯特的错误结论建立在他对从他的理论的广义协变性假设得出的收缩的比安基恒等式的作用的误解基础上。

[122] 13. 外尔指出，“有必要像处理引力现象一样，把电磁现象看作宇宙几何的产物”(1921)。

14. 另见 Schrödinger(1950)。

15. 人们晚得多承认的一些规范不变的量子场论的几何方面，将在 11.3 节讨论。

16. 见 Papapetrou(1949)，Sciama(1958)，Rodichev(1961)，Hayashi and Bregman(1973)，Hehl *et al*(1976)。

17. 拓扑学是数学的一个分支，是定性地处理事物彼此或自身连结时的方式。 拓扑学只关心连接，而不关心形状、大小或曲率。 因此，当几何图形经过连续而光滑的变形而不撕破它成为另一个图形时，它们就是拓扑等价的。 诸如对于奇点问题至关重要的“时空到达尽头或有边界(越过边界，时空停止存在)吗？”问题，以及对于黑洞的形成与存在至关重要的，也对宇宙学至关重要的“时空的哪些区域能够彼此发送信号？ 哪些区域不能彼此发送信号？”问题，都是拓扑学问题。



第二篇

基本相互作用的量子场纲领



[123] 这一篇将分析基本相互作用的量子场纲领(QFP)概念基础的形成，特别关注量子场纲领提出的基本本体论和基本相互作用的传播机制。第6章沿着两条线路来重构1927年以前的量子物理学史，这两条线路分别为：(i)原子系统力学运动的量子化；(ii)波场的量子化。这章也将对海森伯和玻尔(Niels Bohr)刻画量子力学特征时提出的不确定性(uncertainty)和互补性(complementarity)予以描述。第7章以历史的和批判的眼光，对量子场论(QFT)的奠基者和后来的阐释者关于量子场论的概念基础所采取的立场，予以评论。这章的前三节是对量子场论早期历史中出现的本体论转换即从粒子本体论转换到一种新的本体论进行分析。7.4节考察体现在狄拉克真空观念中的量子场论的最初本体论承诺所面临的困境。7.5节是对局域耦合、虚量子交换、不变性原理的观念演化的重构，这些观念被猜想为由量子相互作用所遵守，因此对相互作用的形式强加了限制。7.6节回顾了20世纪40年代晚期、50年代初期对发散的确认和重正化纲领的形成。第8章总结了量子场纲领的本质特征，它的兴衰沉浮和各种探索替代方案的尝试，及其直到20世纪70年代初以规范场论的形式的再次复兴。

第6章 量子理论的兴起

[125] 相对论的起源与探讨相干场理论表述的这一电磁概念的发展密不可分；根据相干场论，所有的电磁行为能够以一种连续的方式变化。相反，量子理论是从原子概念的发展中产生的，这一发展的特征是承认经典物理学中的观念应用于原子现象时有一个根本的限制。这个限制以所谓的量子公设(quantum postulate)表达，即把任何原子过程归因于一个由普朗克作用量子表征的本质上不连续的过程。

量子场论是量子理论概念发展的一个新阶段，前身为旧量子理论和非相对论性量子力学，这两种理论实质上是对原子和辐射之间相互作用的初步分析。本章将重新审查与量子场论的兴起相关的量子物理学的一些特征。

6.1 运动的量子化

在解决物质与辐射之间平衡的问题上，普朗克(1900)指出，热辐射定律在描述原子过程时要求有非连续性成分。根据普朗克的描

述，在由与辐射相互作用的线性谐振子表征的原子的统计行为中，只应考虑其能量是量子 $h\nu$ 的整数倍的振动态，其中 h 是普朗克常量，ν 是谐振子的频率。

普朗克本人认为，能量的不连续性只是原子的一个属性，他不愿意把能量量子化的思想应用于辐射本身。而且，普朗克只赋予这些谐振子以纯形式上的性质；在这些性质之中，缺乏真实原子的一个本质属性，即改变其辐射频率的能力。 [126]

这与其他诸如维恩（Wilhelm Wien）和哈斯（Arthur Haas）等物理学家的思想完全不同。维恩（1909）把电磁谐振子看作真实的原子，它们除了能够吸收和发射辐射能量外，还具有其他特性。例如，它们能被紫外线或X射线电离。根据维恩，“如果能量子从根本上说有任何物理意义的话，那么它只可能从原子的一个普适属性中导出”（1909）。

哈斯（1910a，b，c）通过仔细研读维恩关于原子结构问题的论文，以及J·J·汤姆孙（J. J. Thomson）同一主题的著作《电和物质》（*Electricity and Matter*，1904），从而通向了用真实原子取代普朗克所使用的理想赫兹振子，并将作用量子的本质与原子结构联系起来。通过假设原子的电子势能 e^2/a（a 是氢原子的半径）可以用普朗克能量子 $h\nu$ 描述，即

$$|E_{\text{pot}}| = h\nu, \tag{1}$$

并运用离心力等于库仑引力的经典关系 $m\omega^2 a = e^2/a^2$，哈斯获得了一个方程

$$h = 2\pi e(am)^{1/2}, \tag{2}$$

在方程（1）中，频率 ν 被认为与电子的轨道频率完全一致，$\omega =$

$2\pi\nu$。[1] 作用量子 h 和原子的参量 e、m 和 a（或说原子的量值）之间这一关系的确立，无疑是哈斯对原子运动的量子化思想的概念发展的巨大贡献。

索末菲（Arnold Sommerfeld）不愿意用原子量值来解释 h，而宁愿把原子的存在看作是所有基本作用量子存在的结果。在他看来，作用量子，作为一个新的物理事实，不能从其他事实或原理推出。1911 年在布鲁塞尔举行的第一届索尔韦会议上，索末菲宣称：

> 对我来说，对 h 作电磁的或力学的“说明”，如同对麦克斯韦方程组作力学“说明”一样，是不合道理和没有希望的……几乎不能怀疑，如果物理学需要一个新的基本假说，而这个假说必须要对我们的电磁世界观增加一个新奇的成分，那么，对我来说，作用量子的假说似乎比所有其他假说更好地实现了这一任务。
>
> （Sommerfeld，1911b）

这是他对在同年早些时候提出的量子理论基本假说（1911a）的进一步[127] 发展，据此，电子和原子之间的相互作用确定地和唯一地由普朗克作用量子支配。在评论索末菲的观点时，洛伦兹说：

> 索末菲不否认在普朗克常量 h 和原子的量值之间有一种联系。这能用两种方式表达：或者普朗克常量 h 由这些量值确定（哈斯的观点），或者这些量值（它们被认为属于原子）依赖于普朗克常量 h 的大小。我认为在这些观点之间没有大的不同。
>
> （Lorentz，1911）

玻尔受索末菲上述观点的强烈影响。从1912年3月中旬到7月底，玻尔一直在曼彻斯特的卢瑟福研究所(Rutherford's Institute)工作，他接受了卢瑟福的原子的行星模型。这一模型所面临的严重问题是：一个带正电荷的原子核如何能与一个带负电荷的旋转电子保持平衡？是什么阻止了这个电子跌落进原子的中心？这个关于稳定性的问题实质上构成了玻尔工作的起点。

到1912年年中，玻尔开始确信卢瑟福模型要求的稳定性是非力学起源的，只能由量子假说提供：

> 这个假设是，对于任何稳定的环(任何出现在天然原子中的环)，在环中电子的动能和旋转时间之间有一个确定的比率，(并且对这个假说来说)并不企图给出一个力学的基础(这似乎没有希望)。
>
> (Bohr，1912)

与定态(“稳定环”)相联系的力学运动的量子化规则，可看作玻尔对普朗克关于谐振子可能能量值的最初假设的一个合理推广：它涉及这样一个原子系统，在这个原子系统中，运动的力学方程的解完全是周期的或倍周期的，因此粒子的运动能被表示为离散的谐振动的叠加。

玻尔的伟大成就是把卢瑟福的原子模型和普朗克的量子假说进行了综合，在这个综合中，他引入了一系列假设，包括原子的定态假设，以及原子从一个定态跃迁到另一个定态时发射或吸收辐射的频率的假设。通过运用维恩的结果、德比耶纳(André Debierne)的结果、特别是斯塔克(Johannes Stark)的结果[2]，加上这些假设，玻尔

(1913a，b，c)得以给出了支配元素线谱的主要定律，特别是氢原子光谱的巴耳末公式的简单解释。[128] 玻尔的原子光谱理论使得原子运动的量子化思想具体化，从而可以看作这一思想发展中的一个里程碑。

6.2 辐射的量子化

当普朗克引入量子思想描述纯辐射的光谱性质时，他只是把量子化手段应用于可称重物质，即物质振子。然而，他没有意识到，他的提议暗示着这样一个事实：经典场本身需要有一种新的思想；这种新思想认为，量子是辐射所固有的，它应该被想象为一种自由翱翔的粒子。他的推理声称只涉及对物质和辐射之间的相互作用的修正，因为这一相互作用完全难以理解；但自由电磁辐射不需要修正，因为人们对于自由电磁辐射已有了较好的理解。

与之形成对照的是，当爱因斯坦在 1905 年提出光量子的概念时(1905a)，他敢于挑战高度成功的纯辐射的波动理论。在他“完全革命性”的论文中，爱因斯坦强调说：

> 应该记住，光学观测值是指波随时间流逝的平均值，而不是瞬时值。尽管衍射、反射、折射、色散等理论得到了完全的实验证实，然而可以想象，如果把一个用连续三维函数进行运算的光的理论应用于光的产生和转化现象，将导致与经验的不相容。
>
> (同上)

通过引入辐射的微粒结构，爱因斯坦似乎给自己确立了任务：消除、至少是部分消除横亘在本质上离散的物质原子理论和本质上连续的电磁场论之间的深刻差异。他提出，“光的能量是由有限数量

的、局域在空间不同点上的能量子所组成”，并且这些量子“只能作为单元被产生或吸收”（同上）。显然，光量子假说是关于自由电磁辐射的量子性质的一个断言，它应该被扩展到光和物质的相互作用。这的确是非常革命性的一步。

爱因斯坦在1906年注意到，普朗克辐射理论所依据的两个关键性公式彼此矛盾(1906a)。根据麦克斯韦的电动力学定律，普朗克(1900)通过使用振子模型并心照不宣地假设一个振子的振幅和能量是连续变化的，获得了辐射密度 ρ 和一个振子的平均能 U 之间的关系式

$$\rho = 8\pi\nu^2/c^3 U, \qquad (4)$$ [129]

然而，在推导下面这个表达式

$$U = h\nu/(e^{h\nu/KT} - 1) \qquad (5)$$

（这一表达式决定性地不同于统计热力学中的均分定理 $U = KT$）的过程中，普朗克假设了离散的能量级。其困境在于同时应用了方程(4)和(5)，而这两个方程是从相互矛盾的假设中推导出来的。

爱因斯坦指出，实际上，普朗克的理论除了包含能量的离散假设之外，还包含有第二个假设。那就是，当振子的能量被量子化时，方程(4)必须继续保持，即使它的推导基础（即连续性）已被去除。这第二个假设构成了与经典物理学的进一步分离，表明“普朗克在他的辐射理论中，（心照不宣地）在物理学中引入了一个新的假说性原理——光量子假说”(Einstein, 1906a)，尽管普朗克本人迟至1913年仍拒绝这一假说。

同年，埃伦费斯特(Paul Ehrenfest, 1906)以相同的精神独立地讨论了相同的问题，注意到这一点是有趣的。他指出，如果我

们作出如下假设，即存在于标准频率模式中的场能总和只能是 $h\nu$ 的整数倍，那么，通过对辐射腔的固有振动幅度与物质振子的坐标进行类比，并运用辐射腔共振的瑞利—金斯(Rayleigh-Jeans)求和，有可能会得到普朗克能谱。相同的假设使得德拜(Peter Debye, 1910b)实际上得到了普朗克能谱。参照瑞利—金斯的模密度 $N\mathrm{d}\nu = 8\pi\nu^2/c^3\mathrm{d}\nu$，德拜在赋予每一个自由度以平均能量 $U = h\nu/(e^{h\nu/KT} - 1)$ 之后，直接获得了普朗克的辐射公式。于是，他指出，普朗克定律是从能量的量子化这唯一的假设推出的，"不需要谐振子的中介作用"(1910a)。

1909 年，爱因斯坦给出了支持其光量子假说的新论据(1909a)；这一新论据与他 1905—1906 年间建立在平衡统计分析(依赖于体积内的熵)基础上的论据不同，它建立在关于黑体辐射能量和动量的统计涨落的分析基础之上。他获得的能量和动量涨落公式是：

$$\langle \varepsilon^2 \rangle = (\rho h\nu + c^3\rho^2/8\pi\nu^2)V\mathrm{d}\nu \qquad (6)$$

[130] 以及

$$\langle \Delta^2 \rangle = 1/c(\rho h\nu + c^3\rho^2/8\pi\nu^2)A\tau\mathrm{d}\nu \qquad (7)$$

其中 V 是辐射腔的体积，A 是腔中放置的镜子的面积，τ 是时间间隔。看来似乎有两种独立的产生涨落的原因。第一种机制(独立的光量子)将单独导致维恩定律，第二种机制(经典波)将单独导致瑞利—金斯定律。两种机制中的任何一种都不能单独导致普朗克定律，但两者的结合则可以。

运用这样一种分析，爱因斯坦于 1909 年 9 月 19—25 日在萨尔茨堡会议上宣布：

因此，在我看来，理论物理学发展的下一个阶段将带给我们一个关于光的理论，这个光的理论能被解释为一种波动理论和发射理论的融合……在我们关于光的本质和构成的观点中，一种深刻的改变是绝对必要的。

(1909b)

为了证明这两种结构属性(波的结构和量子的结构)不必是不相容的，爱因斯坦在他论文的结尾提出一个概念，这个概念后来被德布罗意在他1923年的论文(1923a，b)中重新拾起[3]，成为量子力学和量子场论后来发展的基础。这个概念可概括如下：电磁场的能量假定集中在被力场围绕的奇点中，力场遵从叠加原理，因此充当着与麦克斯韦理论的场相似的波场。

1909年爱因斯坦关于光的结构的观念有了一个细微的变化。早期他倾向于用粒子完全取代场的观念。也就是说，粒子是唯一的实在，表观场应该还原为粒子间的直接相互作用。1909年，爱因斯坦仍然钟情于这种光的微粒模型，这一点在他1909年5月23日写给洛伦兹的信中得到证实：

我把光量子假定为由延展的矢量场包围着的点，这个矢量场以某种方式随着距离的增大而减弱。这个点是一个奇点，没有它矢量场不能存在……矢量场应该由所有奇点的运动位置完全决定，这样描绘辐射的参量个数才是有限的。

(1909c)

但是，我们能在同一封信中察觉出一个细微的变化：

然而，对我来说，重要的似乎不是关于这些奇点的假设，而是关于这种场方程的书写；它们有解，根据这些解，有限数量的能量能够在确定的方向上以光速 c 做没有耗散的运动。人们能够设想这一目标可以通过稍稍修正麦克斯韦的理论而实现。

[131]

（同上）

在他的萨尔茨堡论文(1909b)中，场的物理实在也得到承认，这是后来称为波粒二象性(wave-particle duality)显现的第一条线索。[4]

波粒二象性的概念假定：一个物理实体既具有波的实在性，又具有粒子的实在性。然而，完全成熟的关于辐射的粒子概念[5]的一个重要要素仍然缺乏。爱因斯坦在1905年引入的光量子实际上只是能量子。直到1916年，爱因斯坦本人才清楚地提到光量子的动量，尽管方程(7)中的第一项可能已导致他有了这一想法。正是斯塔克(1909)从这一项得出如下陈述："加速电子所发射的总电磁动量不等于零，而是由 $h\nu/c$ 给出。"

爱因斯坦在引力上投入了几年精力之后，由于直接受到洛伦兹发表的关于涨落的计算(1916)的激励——涨落问题是爱因斯坦以前处理过的问题(1909a，b)——于1916年又回到辐射问题。这时，量子理论已焕发出新的面貌。玻尔已在他关于氢原子和氢光谱的理论中开启了量子思想应用的领域。玻尔的工作显著地影响了爱因斯坦的思想。这可从爱因斯坦关于"原子的内部能态"和"从能态 E_m 到能态 E_n 的跃迁是通过吸收或发射一个确定频率 ν 的能量辐射"的新思想中清楚地看到。通过运用这些思想，并引入跃迁概率的新概念和新的系数 A_{mn}、B_{mn} 和 B_{nm}（分别代表自发发射、受激发射和吸收），爱因斯坦重新推导了普朗克定律(1916c)。其时玻尔已假定了他的频

率条件 $E_m - E_n = h\nu$，因而爱因斯坦的工作在普朗克的辐射理论和玻尔的光谱理论之间架起了一座桥梁。

在这篇论文中，爱因斯坦本人认为比上述结果更为重要的，是他对于辐射过程的定向性质的分析所隐含的结论："不存在球面波的辐射。"(1916c)因此，"一个适当的关于辐射的量子理论的建立看来几乎不可避免"(同上)。至此，完全成熟的关于辐射的粒子概念出现了，在这个概念中，一个光量子携带动量 $h\mathbf{K}$(此处 $\mathbf{K}$ 是波矢，$|\mathbf{K}| = \nu/c$)。

[132] 在爱因斯坦光量子思想的影响下，辐射的微粒概念和波动概念之间的冲突变得越来越尖锐。然而，在 1921 年到 1924 年间，人们清楚地认识到，光量子假说不但能应用于斯特藩—玻尔兹曼定律(Stefan-Boltzman law)、维恩位移定律和普朗克定律，而且也能应用于其他光学现象，诸如多普勒效应(Doppter effect；Schrödinger，1922)和夫琅禾费衍射(Fraunhofer diffraction；Duane，1923；Compton，1923a)，这些现象已被看作光的波动概念的无可辩驳的证据。但正是康普顿(Compton)关于 X 射线散射的实验研究给了量子观点以坚定的经验基础。

1923 年，康普顿(1923b)和德拜(1923)各自推导了一个光量子在一个静止电子上散射的相对论性运动学方程：

$$h\mathbf{K} = \mathbf{P} + h\mathbf{K}', \quad hc\,|\mathbf{K}| + mc^2 = hc\,|\mathbf{K}'| + (c^2\mathbf{P}^2 + m^2c^4)^{1/2} \tag{8}$$

其中 $h\mathbf{K}$、$h\mathbf{K}'$ 和 $\mathbf{P}$ 分别代表光量子的初动量、末动量和电子的末动量。这些方程隐含初始光量子和终末光量子的波长差 $\Delta\lambda$ 为：

$$\Delta\lambda = (h/mc)(1 - \cos\theta), \tag{9}$$

其中 θ 是光量子的散射角。康普顿发现这个关系式在误差许可的范围内得以满足，他得出结论说："这个理论的这一实验支持非常令人信服地表明，辐射量子携带动量和能量。"(1923b)

康普顿的贡献在于使物理学界的大多数成员接受了辐射的量子观点，这可与一个世纪以前菲涅耳使物理学界的大多数成员接受了光的经典波动理论的贡献相媲美。然而，爱因斯坦对康普顿实验的反应却是有趣地谨慎：

> 康普顿实验的肯定结果证明，不但关于能量传递，而且关于动量传递，辐射行为似乎都是由离散的能量发射物组成的。
>
> (Einstein，1924)

在爱因斯坦开创的光的量子概念取得胜利的情况下，更能阐明爱因斯坦立场的是他对光的波粒二象性的清晰阐述："因此，现在有两种光的理论，两者既必不可少，又没有任何逻辑关联。"(同上)

爱因斯坦的立场可被看作反映了这样的事实，即不能在两种相互[133]竞争的概念之间硬要作出非此即彼的抉择。对解释涉及光和物质之间相互作用的光学过程来说，量子观点似乎必不可少，而像干涉和衍射现象似乎需要光的经典波动说的概念工具。甚至更为糟糕的是，由无可争辩的实验证据支持的光量子假说，只有通过运用它自己的反面，即波动假说，才会变得物理上有意义。正如玻尔、克拉默斯和斯莱特(Bohr，Hendric Kramers，and John Clarke Slater，1924)指出的，这是因为光量子是由只有通过应用诸如衍射的波动概念才能测量的频率所定义的。

6.3 矩阵力学的诞生

在 1919—1925 年间，原子物理学中的研究工作主要是基于玻尔的理论。玻尔在 1913 年的三篇论文中(1913a，b，c)，把他的基本观点表达如下：(i)由量子条件确定的定态，(ii)频率条件 $E_1 - E_2 = h\nu$，它显示谱线的频率各自与两个态相关联。此外，还有一个重要的启发性原理，这一原理的推广形式后来被玻尔称为对应原理(correspondence principle)，它在 1918 年被清楚地表述为：

> 在相互之间差别极其微小的相继定态中的运动的极限情况下，[由频率定律计算的频率，]将倾向于与根据通常的辐射(来自处于定态中的系统的运动)理论所预期的频率相一致。
>
> (1918)

这个原理使得在原子中保留电子运动的经典描述成为可能，但同时允许对结果作出某种修正以符合观测数据。

然而，仍然留有两个困难。其一，这个理论是经典理论和量子假说的一个糟糕的大杂烩，缺乏逻辑一致性。其二，根据定态的力学模型，量子条件很容易与原子中电子的周期轨道相联系，并且谱线的光学频率应该与电子运动的傅里叶轨道频率相一致，这是一个从未被实验证实的结果，在这个结果中，所观测谱线的频率总是与两个轨道频率的差相关联。

在爱因斯坦关于跃迁(在此，跃迁被定义为与一个原子的两个态有关的量)的论文(1917b)影响下，物理学家的注意力从定态的能量移向定态之间的跃迁概率。正是克拉默斯(1924)开始认真地研究了原子的散射问题，并认真地在辐射条件下的玻尔模型的行为与爱因斯坦 [134]

系数 A_{mn}、B_{mn} 和 B_{nm} 之间建立了联系。

玻恩(1924)扩展了克拉默斯的观点和方法，把它们应用于辐射场与一个辐射电子之间的相互作用以及一个原子的几个电子之间相互作用的情形。玻恩在执行这一[对应原理]纲领时表明，如果某种微分能用相应的差分替代，就能实现从经典力学到他称为“量子力学”的转变。对于大 n 和小 τ 来说，根据玻尔的对应原理，从定态 $n' = n - \tau$ 发出的跃迁的量子理论频率 $\nu_{n,\,n-\tau}$ 与经典频率 $\nu(n,\tau)$ 相一致。也就是说，在态 n 中，经典运动的基本频率的第 τ 个谐波为：

$$\nu_{n,\,n-\tau} = \nu(n,\ \tau) = \tau\nu(n,\ 1), \tag{10}$$

其中 $\nu(n, 1)$ 是经典基本频率，等于哈密顿量对作用量的导数：$\nu = \mathrm{d}H/\mathrm{d}J$。[6]在这种情况下，将(经典频率) $\nu(n,\ \tau) = \tau\mathrm{d}H/\mathrm{d}J = (\tau/h)\mathrm{d}H/\mathrm{d}n$ 与(量子频率) $\nu_{n,\,n-\tau} = \{H(nh) - H[(n-\tau)h]\}/h$ 进行比较，玻恩认为，通过用差分 $H(n) - H(n-\tau)$ 代替微分 $\tau\mathrm{d}H/\mathrm{d}n$，就能从 $\nu(n,\ \tau)$ 获得 $\nu_{n,\,n-\tau}$。

克拉默斯和海森伯(Kramers and Werner Heisenberg, 1925)讨论了散射光频率不同于入射光频率的散射现象。他们的方法明显与玻恩的方法(1924)有关联，完全按照与两个态相关联的量来实施，使用多重傅里叶级数，并用差商取代微商。这里，散射光量子不同于入射光量子，因为在散射过程中原子发生跃迁。当他们试图写下这些情形中的色射公式时，他们不但不得不论及爱因斯坦的跃迁概率，而且不得不论及跃迁幅度，并且不得不让两个幅度相乘，比如说从态 m 到态 n_i 的幅度乘以从态 n_i 到态 k 的幅度，然后对整个中间态 n_i 求和。

这些乘积的和几乎全是矩阵的乘积。从这些矩阵的乘积到用相

应的矩阵元取代电子轨道的傅里叶分量，只是很小的一步，海森伯在其历史性论文《关于运动学和力学关系的一个量子理论转译》(*On a quantum theoretical re-interpretation of kinematic and mechanical relations*, 1925)中实现了这一步。

海森伯论文中的主要观点是：第一，在原子范围内，经典力学不再有效；第二，他寻找的是一种必须满足玻尔对应原理的新力学。[135] 关于第一点，海森伯(1925)写道：

> 爱因斯坦—玻尔的频率条件已显示了与经典力学的这样一种彻底分离，更确切地说，是与作为这一力学基础的运动学的彻底分离，甚至对于最简单的量子理论问题，经典力学的有效性也完全不能保留。

在寻找新的运动学时，他不得不抛弃把量 x 解释为一个赖时位置的运动学解释。那么，在[新的]运动方程中，何种量将取代 x 呢？海森伯的观点是引入依赖于两个量子态 n 和 m 的"过渡量"。例如，在周期运动的经典情形中，$x(t)$的傅里叶展开式为

$$x(t) = \sum_{\alpha=-\infty}^{\infty} a_\alpha e^{i\alpha\omega t}, \tag{11}$$

他写下了新的一项 $a(n, n-\alpha)e^{i\omega(n, n-\alpha)t}$，以取代 $a_\alpha e^{i\alpha\omega t}$ 项。海森伯宣称，强度，因此也就是 $|a(n, n-\alpha)|^2$，与不可观测的经典函数 $x(t)$形成对照，是可观测量，由此推动了 $a(n, n-\alpha)$ 的引入(同上)。

在这篇论文中，海森伯特别强调的一个思想是关于"建立一个完全基于可观测量间关系之上的量子理论力学"的假设(同上)。这个建议曾作为量子力学成功的根基而被大加赞扬。然而，事实是，一

方面，海森伯把电子的位置看作不可观测的，但他错了，因为根据充分发展的量子力学，一个电子的三个坐标 x，y 和 z 是可观测的(Van der Waerden，1967)。另一方面，薛定谔的波函数却是不可观测的，但没有人怀疑它的理论意义(Born，1949)。因此，在爱因斯坦和其他人的批评下，海森伯在晚年放弃了他的主张(Heisenberg，1971)。

关于第二点，海森伯已做了三个观测结论。第一，一个满足对应原理的重要策略是玻恩用差商取代微商的策略。第二，力学的哈密顿形式

$$Q_r = \partial H/\partial P_r,\ P_r = -\partial H/\partial Q_r, \tag{12}$$

能通过用相应的矩阵取代所有动力学变量得以保持。由矩阵表示物理量可以被看作新量子力学的本质；并且矩阵的引入，比起关于可观测量的主张来，更可以看作海森伯对量子理论概念发展的重要贡献。第三，运用这一表示法，旧量子条件 $\oint P\mathrm{d}Q = nh$ 能被改写为对易关系式：

[136]

$$PQ - QP = (h/2\pi\mathrm{i})I \tag{13}$$

作为原子运动量子化的一个特殊方案，矩阵力学的最初形式显然只适用于具有离散能级的封闭系统，但不适用于自由粒子和碰撞问题。针对量子理论的这些应用，需要一些新的概念工具。

6.4 物质的二象性、个体性和量子统计

德布罗意认为，爱因斯坦关于辐射的波粒二象性的观念具有绝对普适性，有必要扩展至所有的物理世界。德布罗意划时代的新原理："任何移动的物体可能伴随一个波，不可能使物体的运动和波的传播分开"，首先在他 1923 年的一篇论文(1923a)中阐述。在 1923 年写的另一篇论文中，德布罗意指出，一束电子流穿过一个尺寸比电

子波的波长小的孔径，“将显示衍射现象”（1923b）。

德布罗意的主要观点是优美的，并且是对爱因斯坦的辐射工作的精确完善。在爱因斯坦把粒子性赋予辐射的地方，德布罗意把波的性质赋予物质。通过以下关系式，德布罗意把物质波的频率 ν 和波长 λ 与粒子的能量 E 和动量 P 联系起来：

$$E = h\nu, P = h/\lambda \tag{14}$$

物质波的概念，的确使德布罗意能给出旧量子条件 $\oint P\mathrm{d}Q = nh$ 一个很神奇的几何解释：既然 $P = h/\lambda$ 暗示 $\oint 1/\lambda \mathrm{d}Q = 2\pi\tau/\lambda = n$，那么量子条件刚好就是经过轨道周长的波长总数是一个整数的条件。

在康普顿实验最终让许多物理学家相信光量子的实在性之际，爱因斯坦站在了德布罗意一边；德布罗意提议相同的二象性不仅为辐射所有，而且一定为可称重物质所有。在 1925 年关于理想气体的量子理论的论文中，爱因斯坦根据对于涨落的分析，为支持德布罗意的观点提供了新论据。

爱因斯坦的一个出发点是玻色(Satyendra Bose)的工作。玻色(1924)只运用统计力学的方法，避免对经典电动力学的任何参考，就直接从爱因斯坦的光量子假说推出了普朗克定律。他把量子看作粒子(1924)。但正如埃伦费斯特(1911)已经意识到的，具有显著个体性的独立量子只能导致维恩定律，而不能导致普朗克定律。因此，如果光量子被用来解释普朗克分布，那么它们必然缺乏统计独立性（通常与自由粒子相联系），而显示某种关联[正如物理学家通常做的，这或者能被解释为这种现象的波动特征的一个迹象；或者正如爱因斯坦做的(1925a)，能被解释为粒子处在一种很神奇的影响下的一 [137]

个迹象]，并遵守一种与独立粒子的玻尔兹曼统计不同的统计，即遵守不可分辨粒子的统计；这一点首先由埃伦费斯特及其合作者于1915年发展起来(1915)，尔后又为玻色于1924年重新独立发现(1924)。通过运用一种不寻常的后来称作玻色—爱因斯坦统计[8]的计算方法，玻色心照不宣地把这些关联并入他的理论，有效地否定了光量子的个体性。

根据“辐射和气体间有一种深远的形式关系的假设”，爱因斯坦(1925a)通过让 $V\rho d\nu = n(\nu)h\nu$，$\langle \varepsilon^2 \rangle = \Delta(\nu)^2(h\nu)^2$ 和 Z(每个间隔中态的数目) $= (8\pi\nu^2/c^3)Vd\nu$，把他的电磁辐射的均方能量涨落的公式(方程(6))改写为：

$$\Delta(\nu)^2 = n(\nu) + n(\nu)^2/Z(\nu) \tag{15}$$

然后，他表明方程(15)对于他的量子气体同样有效，只要 ν 在后一种情形中用 $\nu = E/h = P^2/2mh$ 来定义，并且使用玻色统计而不是玻尔兹曼统计。方程(15)中第一项是唯一代表独立粒子气体的项，第二项对应于辐射情形中经典波干涉项。虽然第一项对于辐射来说是让人感到奇怪的项，但是对气体情形来说，问题是怎么处理第二项，它体现了粒子的不可分辨效应。由于第二项与辐射情形中的波相联系，因此导致爱因斯坦“通过把气体与辐射现象(即波动现象)联系起来，用解释气体的相应方式去解释它”(同上)。

但是，这些波是什么？爱因斯坦建议，一个德布罗意类型的波场应该与气体相联系，他寻求德布罗意波的含义：

> 这个物理性质暂时仍然模糊的波场，原则上必定要用与之相应的衍射现象来说明。因此，穿过孔径的一束气体分子，必定经历

[138]

一个类似于光线的衍射。

(同上)

然而，他很快补充说对于可控孔径，这种效应极其微弱。

受爱因斯坦论文的激励，埃尔萨瑟(Walter Elsasser，1925)提出，能量值低于 25 eV 的慢电子将理想地适合于用来检验“[这个]假设，即对于粒子的每一个平移运动，必定与决定粒子的运动学的波场相联系”。他还指出，拉姆绍尔(Carl Ramsauer)、戴维森(Clinton Davisson)和孔斯曼(Charles Kunsman)的现有实验结果，似乎已给出物质波的衍射和干涉的证据。

6.5 波动力学的诞生

受德布罗意的思想和“爱因斯坦首先给出的具有远见的评论[Einstein，1925]”的启发，薛定谔(Erwin Schrödinger)在他《论爱因斯坦的气体理论》(*On Einstein's gas theory*，1926a)一文中，把德布罗意的观点应用于气体理论。根据薛定谔，爱因斯坦的新气体理论的基本要点是：将所谓的玻色—爱因斯坦统计应用于气体分子运动。考虑到如下的事实，即除了玻色对普朗克定律的推导，还有德拜的推导，是通过把“自然”统计应用于场振子或辐射自由度而实现的外，爱因斯坦的气体理论也能通过把自然统计应用于表征分子的波场振子而获得。于是，薛定谔声称：“这意味着，除了认真地接受德布罗意—爱因斯坦关于运动粒子的波动说之外，我们别无选择。”(1926b)

在 1926 年的另一篇论文(1926c)中，薛定谔发展了德布罗意在经典力学和几何光学(此两者都不能适用于小尺度情况)之间的类比，在

波动力学和波动光学之间进行了类比。他由此主张：

> 我们必须按照波动说来严格地看待物质，也就是说，为了形成［物理过程的］微观结构的图景，我们必须从波动方程出发，而不是从基本的力学方程出发。
>
> （同上）

他还指出，关于微观结构的新近的非经典理论，

[139]

> 与哈密顿方程及其解的理论［有］一个非常密切的关系，也就是说，与已经最清楚地指出力学过程具有真实的波动特征的经典力学的形式有一个非常密切的关系。
>
> （同上）

运用哈密顿原理，薛定谔用一个变分问题替代了量子条件。也就是说，他试图找到这样一个函数 ψ：对于它的任一变化，“对整个坐标空间的”哈密顿密度的积分，“是稳定的，ψ 总是实数，单值的、有限的，并且直至二阶都是连续可微的”(1926b)。在这个程序的帮助下，

> 量子能级立即作为波动方程的本征值确定下来，波动方程本身带有其自然边界条件，［并且］在如下两个分立的步骤中不再出现：(1)所有路径在动力学上都可能有定义；(2)丢弃那些解的大部分，通过特殊的假设［量子条件］选择不多的一些解。
>
> （同上）

对薛定谔引入的波函数的解释，是量子物理学史上最困难的问题，这个问题将在7.1节给予考察。

6.6 不确定性与互补性

量子力学[9]的兴起已提出了如何刻画它的特征的困难问题。人们普遍接受，量子力学“的特征是承认经典物理学中的观念应用于原子现象时有一个根本的限制”(Bohr，1927)。这个限制的理由，正如海森伯解释的，被认为在于如下这一事实：“电子和原子”拥有的物理实在性，根本不同于经典实体拥有的实在性(Heisenberg，1926b)。但是，如何描述这一不同，即使在接受这个根本差别的量子理论学家中间，也是一个有争议的话题；这种争论非常不同于他们与有经典取向的物理学家，诸如爱因斯坦、薛定谔和德布罗意之间的争论。

量子理论学家群体的普遍策略是：首先，保留而不是抛弃经典物理学的概念；其次，通过禁止经典物理学和量子物理学概念的同时充分使用，或者更精确地说，通过把它们分为不相交的两类(这种分类的规则是只有属于同一类的概念才能同时应用于假定由量子力学描述的实体和过程)，给它们强加限制。玻尔是为经典概念的保留在哲学 [140] 上作辩护的主要人物。他认为，人类的概念化能力脱离不开经典的直觉概念的限制。因此，假若没有这些经典概念，主体间关于经验证据的无歧义的交流就不可能实现。

当然，有趣的是这些限制的具体形式。在这点上的概念发展，人们认为始于玻恩关于碰撞的研究(1926)；在这一研究中，玻恩提出了概率解释。然而，论文的内容要比这丰富得多。既然在玻恩的散射理论中，非对角矩阵元使用由傅里叶变换描述的波函数计算，因此泡利(Pauli)看出，它必定包含如下结论：

一个人能用 p 眼光观察世界，也能用 q 眼光观察世界，但如果同时用两种眼光观察世界，就会误入歧途。

(1926)

玻恩论文的这一推论为狄拉克(1926b)用更高的数学清晰性所阐明。狄拉克在他关于变换理论的工作中表明，玻恩的概率幅实际上是不同(正则共轭)表象间的一个变换函数。考虑到一对正则共轭变量受诸如 $PQ - QP = (h/2\pi i)I$ 的对易关系约束这一事实，狄拉克的理论给海森伯提供了一个用于将玻恩的概率幅中包含的关于共轭变量的同时可测性的限制概念化的数学框架，虽然海森伯已经认识到，作为对易关系的一个推论，不可能在同一实验中测量正则共轭变量。海森伯在没有看到狄拉克的论文之前，在 1926 年 10 月 28 日写给泡利的一封信中声称：

因而，在波的表象中，方程 $pq - qp = h/2\pi i$ 总是对应于这样的事实：不可能及时地(或在一个很短时间间隔内)谈论一个确定位置上的一个单色波……类似地，不可能谈论一个有确定速度的粒子的位置。

(转引自 Hendry，1984)

刻画量子力学特征的下一个步骤，是海森伯关于不确定关系的论文。在这篇论文中，海森伯(1927)设法建立"关于[电子的]位置、速度、能量等词的确切定义"。他声称"不可能用通常的运动学上的术语来解释量子力学"，因为与这些经典概念不可分割地联系在一起的这些术语是矛盾的，并充满着"关于非连续统理论和连续统理

论、粒子和波的观点的斗争”。根据海森伯，这些直觉概念的修正 [141] 必须由量子力学的数学形式体系来指引，其中最重要的部分是对易关系。他强调，“从量子力学的基本方程，似乎可直接得出对于运动学和力学概念进行修正的必要性”。这里他再次谈到对易关系，并从中获得重要的教益。对易关系不但使我们意识到，我们“有正当的理由来怀疑对于‘位置’和‘动量’这些术语的不加批判的应用”，而且它们也给这些直觉概念施加了限制。

为了推导出关于诸如位置、动量和能态等经典概念的限制的具体形式，海森伯从一种操作主义立场[10]（据此，概念的意义只能由实验测量来定义）出发，并且诉诸思想实验。一个著名的思想实验是 γ 射线显微镜实验，这一实验阐明了位置和动量之间的不确定关系：

$$\Delta P \Delta Q \sim h \tag{16}$$

其中 ΔP 和 ΔQ 是动量和位置的测定误差。不确定关系(16)式意味着，对于物理事件或过程的空间（或时间）中位置的精确描述，就排除了对于伴随那一事件或过程交换的动量（或能量）的精确说明。物理学家在解释他们探索量子领域过程中遇到的奇异特征（诸如量子涨落）时，以不同的方式利用了不确定关系的这一推论。

以狄拉克按照主轴变换来解释的广义变换理论为基础，人们给出了一个与(16)式相似的关于不确定关系的更严格的推导。在一个沿着主轴的参考系中，与一个动力学量（或一个可观测量）相联系的一个矩阵（或一个算符）是对角化的。在一个物理系统上完成的每个实验确定了某个方向，它可能沿着主轴，也可能不沿着主轴。如果它不沿着主轴方向，那么存在一个由主轴的变换公式表示的偶然误差或不准确量。例如，测量一个系统的能量就会使系统进入一个态，在这

个态中，位置有一个概率分布。这一分布由变换矩阵给出，能被解释为两个主轴间夹角的余弦。

海森伯声称，从物理上讲，不确定关系需要每个实验以一种独一无二的方式，把物理量分为“已知量和未知量(或者说成：较精确或较不精确的已知变量)”，并且从实现已知量和未知量的不同区分的两个实验中，得到的结果之间的关系只可能是一个统计的关系。[142] 通过强调测量粒子的实验装置的认知意义，海森伯设法解释并消除掉了量子力学中的特殊项，即干涉项，通常把它解释为量子实体的波动性的一个说明。

海森伯用不确定关系来刻画量子力学，受到玻尔的批评，原因是海森伯忽略了物质和光的波粒二象性。由于普朗克常量很小，在经典领域，二象性不在我们关注的视野之内(在经典领域，我们通常使用的因果时空描述是适当的)。但在原子领域，玻尔(1928)强调，普朗克常量以一种“完全不同于经典理论”的方式把测量仪器与所研究的系统连结在一起。也就是说，“对于原子现象的任何观测将涉及一种与观测介质的相互作用(由普朗克常量刻画)，这种相互作用不可忽略”。因此，玻尔主张，必须认真考虑二象性的概念，只能把一个在时空中不受干扰的系统的观念看作一个抽象概念，“必须推广描述的经典模式”(同上)。

1927 年 9 月 16 日在意大利科莫召开的国际物理学会议上，玻尔基于这一考虑第一次提出了他关于量子物理学的互补性观点。他主张，按照波粒二象性的要求，必须对经典概念进行修正。因为根据经典概念，物理实体只能或者被描述为连续的波，或者被描述为不连续的粒子，而不能同时被描述为两者。在原子领域，玻尔强调，光和物质的波和粒子模式既非矛盾也非佯谬，更确切地说是互补的，即

只有在极端的情形下，才相互排斥。（因此，根据玻尔，海森伯的不确定关系只是互补性的一种特殊情形，因为正则共轭量并不真的相互排斥。）然而，为了完整地描述原子现象，波和粒子这两种模式都是需要的。[11]

1927 年，随着不确定关系和互补原理的系统表述，量子力学除了获得数学形式体系外，也获得了一个解释框架，变成了一门成熟的物理学理论。尽管从那时起就有许多令人困惑的问题和不一致持续至今，诸如波包坍缩(reduction)和量子过程的不可分离性，然而这些困惑和不一致对于我们讨论与量子力学的扩展相关的相对论性量子场论的概念发展没有直接的关联，因此不在本书考虑的范围之内。

注释

[143]

1. 方程(1)应用于氢原子的基态时，导致了与玻尔条件相一致的结果。因此，这里 a 可看作是氢原子的“玻尔半径”的第一个形式。

2. 维恩的工作是在普朗克振子的不同态和真实原子的不同态之间作了一个类比，使得同一个原子的不同能态的概念成为必需。德比耶纳关于原子内部自由度的假说基于辐射现象；斯塔克的光谱起源的观念类似于如下观点：单个价电子产生所有谱线，这些谱线是在电子从一个几乎完全分离的态到最小势能态的连续跃迁过程中辐射出来的。见 Pais (1986)。

3. 德布罗意在他的双解理论中甚至更忠实地继承了爱因斯坦的这个概念(de Broglie, 1926, 1927a, b)。然而，由于玻恩对波函数的概率解释的统治地位，德布罗意新理论的出现，在历史上并没有起什么作用。（见 7.1 节。）

4. 更有甚者，在爱因斯坦关于粒子作为场的奇点的谈话中，他会有意无意地转到这样的思想路线上来：粒子(光量子或电子)被解释为奇点，或非线性场中其他内聚性的结构。这种追求能在爱因斯坦的统一场论和德布罗意的非线性波动力学中找到痕迹(de Broglie, 1962)。

5. 光的量子概念由 G・刘易斯(G. Lewis)命名为光子(Lewis, 1926)。光子除了具有能量 $E(=h\nu)$ 之外，还有动量 $\boldsymbol{P}(=h\boldsymbol{K}$，$\boldsymbol{K}$ 是波矢)，它满足色散定律 $E=c\,|\boldsymbol{P}|$。

6. 由于 H 是 $J(=nh)$ 的函数，因此玻尔频率条件规定 $\nu_{n,\,n-\tau}=\{H(nh)-H[(n-\tau)h]\}/h=\tau\mathrm{d}H/\mathrm{d}J=\tau\nu$。

7. 强度与爱因斯坦发射概率成正比，后者与 $|a(n,\,n-\alpha)|^2$ 成正比。

8. 玻色—爱因斯坦统计适用于一种不可分辨粒子(玻色子)，它们在基本存储单元中的数量不受限定。对于另一种不同的不可分辨粒子(费米子)来说，它们在基本存储单元中的数量只能是 0 或 1，遵循一种不同的量子统计，即所谓的费米—狄拉克统计。

9. 薛定谔表明，在矩阵力学的形式中，或者在波动力学的形式中，它们在数学上和描述上几乎等价(Schrödinger, 1926e)。

10. 同一年，美国物理学家布里奇曼(P. W. Bridgeman, 1927)发表了他的操作主义。

11. 玻尔用其互补原理刻画量子力学。除了按照波和粒子进行互补描述外，还有时空和因果性(在能量和动量守恒的意义上)的互补描述："因此，我们或者有时空描述，或者有运用能量和动量守恒定律的描述。它们是互补的。我们不能同时使用它们。如果我们想运用时空观念，我们必须有外在于和独立于所考虑物体的钟表和测量杆，在那种意义上，我们必须忽略物体和所使用的测量杆之间的相互作用。为了应用时空观点，我们必然受制于通过仪器的总动量的确定。"(Bohr, 1928)

第7章
量子场论概念基础的形成

对量子场论(QFT)的分析，可以根据它的数学结构，其物理描述 [144] 的概念体系，或者它的基本本体。这种分析既可以从逻辑上进行，也可以从历史上进行。这一章只论述与其基本本体相关的量子场论概念基础的起源，而不讨论它的数学结构或它的认识论基础。之所以要讨论一些概念问题，诸如与概率和测量相关的概念问题，仅仅是因为它们与量子场论的基本本体相关，而不是因为它们内在的哲学重要性。这里，我用基本本体表示我们发明量子场论来描述的不可还原的实体。[1] 在物理世界的事实中，经常提到的量子场论基本本体的候选者，是离散的粒子和连续的场。最近也已提出了另一个可能的候选者(时空点)(Redhead，1983)。由于这一章的目的是分析量子场论的概念基础如何奠定的历史过程，而不是今天哲学家论述的量子场论的逻辑结构，因此将不讨论最近提出的那些可能候选者。

这一章的内容，粗略地以年代顺序排列，涉及波函数、量子化、场的量子化、真空、场之间的相互作用，以及重正化等概念的形成和

解释。前两个主题将联系场的量子化来讨论，它们的作用被看作量子场论的概念发展的起点。关于相互作用，它是场论(经典理论和量子理论)的起源和这本书的主题，除了这里给出的简要论述外，进一步的讨论将在第三篇给出。

[145]

7.1 波函数的解释

薛定谔在其波动力学中引入的波函数，正如我们将在7.3节中看到的，被看作一个量子化的场。因此，波函数的解释对于量子场论的本体的讨论是至关重要的，因此也是这第一节的主题。

对于波函数的各种解释的一个总评论，能在雅默(Max Jammer)的《量子力学哲学》(*The Philosophy of Quantum Mechanics*, 1974)一书中找到。我们撇开形而上学的猜测和不成功的物理尝试(诸如隐变量假说，或流体动力学解释和随机解释)，把注意力集中于在量子物理学历史发展中充当引导者的解释上。通常认为，人们已接受的关于波函数的解释是玻恩(1926)首先提出的概率解释。这已被评价为量子力学解释中“决定性的转折点”和“最终阐明物理学解释的决定性步骤”(Ludwig, 1968)。然而，这些断言只是相对于非相对论性量子力学为真，而相对于量子场论则为假。就量子场论而言，情形要比上述观点的持有者想象的更为复杂，因为在全面的概率框架中可以清楚地分辨出一个确定的实在要素。因此，为了下面讨论之便，更为适合的做法是：以一种历史的和比较的方式，对实在论解释和概率解释作一简要介绍。

正如我们在上一章所了解到的，薛定谔的波函数，满足某种波动方程和自然边界条件，它起源于德布罗意的物质波(matter wave)思想。在德布罗意的理论中，与任何物质粒子的运动相联系的是一组

平面波，每一个平面波由一个波函数 $\Psi_{dB} = \exp[i/h(Et - Px)]$ 表示(其中 E 表示粒子的能量，P 表示粒子的动量)，这样，粒子的速度等于波的群速 V_g，波的相速等于 c^2/V_g。用一个波函数 $\Psi_1 = \exp[i/h(Et - Px)]$ 表示频率为 ν 的普通光波(其中 E 表示相应光子的能量 $h\nu$，P 表示相应光子的动量 $h\nu/c$)。比较 Ψ_{dB} 和 Ψ_1，我们注意到，普通光波只不过是附属于相伴光量子的德布罗意波。从这一事实，我们能获得两个不同的本体论含义的推论。一个是实在论倾向的推论，最初由德布罗意本人得出，随后为薛定谔接受；这一推论断言德布罗意波，恰如普通光波一样，是真实的、三维的、连续的物质波。另一个推论本质上是概率的，首先由玻恩得出，然后为许多 [146] 人接受，我将对此作简短的讨论。

这里，我们应该非常仔细地论述用波函数表达的德布罗意物质波思想。为了使这一思想的意思更为清楚，避免在文献中频繁出现的混淆，有必要区分“物质波”这一术语的两种基本用法，这对于阐明量子场论的本体是至关重要的。我们可用两个对照来说明这两种用法：(1)物质波对光波，(2)(物质或光的)物质(实体)波对(物质或光的)概率波。在这个术语的第一种用法中，德布罗意—薛定谔波与物质粒子的运动相联系。也就是说，它与光量子毫无关系。这是德布罗意相波的最初含意，这一点没有疑问。

在文献中令人混淆和引发争议的是这个术语的第二种用法。人们可能认为在玻恩的概率解释发表以后，争论已解决。然而在量子场论的语境中这不是真的。一些通常接受玻恩观点的量子场论物理学家，当他们实际上把波函数看作物质波时，仍然把物质的意义归因于德布罗意—薛定谔波。我将在 7.3 节讨论这种混淆的原因和含义。

薛定谔受德布罗意相波思想的影响，并推广了它。他用一个波函数表征单原子体系，认为波是原子过程的载体。他把波函数"与原子中的某种振动过程"相联系(Schrödinger，1926b)，并把波看作"一个只依赖时间的简谐振动的叠加，它的频率精确地与微观力学系统的光谱项频率相一致"(1926d)。但接着产生了一个大问题：它是什么波？

在1926年的另一篇论文(1926e)中，薛定谔试图赋予波函数以电动力学意义："电荷的空间密度由 $\Psi\partial\Psi^{-}/\partial t'$的实数部分给出。"但是，这里的困难在于，当对全空间求积分时，这个密度等于零而不是一个不依赖于时间的有限值。为了解决这一矛盾，薛定谔在他1926年的另一篇论文(1926f)中，用"这一系统的位形空间中的权重函数" $\Psi\Psi^{-}$ 乘以总电荷 e，来代替电荷密度的表达式 $\Psi\partial\Psi^{-}/\partial t$。这相当于下面的解释：

> (原子)系统的波动力学位形，是许多——严格地讲是全部——运动学上可能的质点力学位形的一个叠加。因此，每个质点力学位形对真实的波动力学位形贡献某种用 $\Psi\Psi^{-}$ 精确表示的权重。
>
> (Schrödinger，1926f)

[147]

也就是说，原子系统"同时存在于所有运动学上可能的位置，而不是'同等地'处于所有的位置"(同上)。薛定谔在这一解释中如此强调：

> Ψ 函数本身不能也不可能直接用三维空间来解释——无论

在这一点上单电子问题是如何经常地倾向于误导我们——因为 Ψ 一般是($3n$ 维)位形空间的函数,而不是实空间的函数。

(同上)

尽管如此,他仍然坚持实在论解释:

但是,在现在的概念背后也有可触知的某种真实存在,即关于电子空间密度的完全真实的电动力学有效涨落。Ψ 函数要做的刚好是允许[由一个波动方程]控制和检查所有这些涨落。

(同上)

因此,他把这种波看作具有某种与电磁波拥有的相同类型的实在性,有一个连续的能量密度和动量密度。

速度为 V 的粒子能用以群速 V 运动的波包来表示;基于这一假设,薛定谔(1926c)建议彻底抛弃粒子本体,主张物理实在由波组成,且仅由波组成。在1961年的一篇论文中,薛定谔把他拒绝粒子作为一个有确定特性或个性的、定义明确的不变客体的理由总结如下:(1)"由于质量和能量的同一性,我们必须把粒子本身看作普朗克能量子";(2)在处理相同类型的两个或多个粒子过程中,"我们必定影响它们的同一性,否则,结果将会完全不真实,也与经验不一致";(3)"根据[不确定性原理],不可能以无可置疑的确定性两次观测到同一个粒子"。

把连续场看作量子力学的基本本体的一个相似但更为精致的观点,由德布罗意(1926,1927a,b)在他的双解理论中提出。德布罗意和薛定谔观点的主要不同在于,薛定谔把物理实在看作仅由波组成,

而德布罗意除了接受波的实在性之外，还接受经典粒子的实在性，而且德布罗意还把粒子看作由延展线性波 Ψ 引导的非线性波在一个奇[148]异区的能量集中。 因此，根据德布罗意，物理实在由波和粒子组成，尽管后者只是前者的一种显现。

波函数的这种实在论解释不得不面对严重的困难。 第一，波不能被看作一个真实的波，因为，正如玻尔在他 1927 年的论文中所指出的，波的相速是 c^2/V_g(V_g 是群速)，它通常大于光速 c。 第二，就波函数 Ψ 的位形空间的维数而言，正如洛伦兹 1926 年 5 月 27 日写给薛定谔的信中所指出的，当波与借助于若干粒子作经典描述的过程相联系时，波就不能从物理上得以解释。 因为那时波变成了一个由 $3n$(n 是粒子数)个位置坐标组成的函数，并需要有一个 $3n$ 维空间表示它。 第三，薛定谔的波函数是一个复函数。 第四，Ψ 函数依赖于表象。 最后的也是最为严重的困难与所谓的波包扩散相关。 洛伦兹在上面提到的信中指出，“波包从来不能长久地呆在一起，保持限制在一个小的体积之中”，因此把波包表示成粒子(“一个相当持久不变的个体存在”)是不合适的。 薛定谔受到在谐振子情形中获得不扩散波包的成功的鼓舞，认为在通常的情形中得到它“仅仅是一个计算技巧的问题”。 但是，海森伯(1927)不久意识到这是一个不可能完成的任务。

现在我转向玻恩的概率解释，这一解释与把粒子看作是量子物理学的基本本体的观点相一致，并迅速变成了一个占支配地位的解释。

玻恩于 1954 年荣获诺贝尔物理学奖，“由于他在量子力学中的基本工作，特别是他对波函数的统计解释”(1954 年 11 月 3 日瑞典皇家学院官方宣布辞)。 然而，玻恩的观点起源于爱因斯坦关于波场与

光量子关系的意见(1909a, b)。在他关于碰撞过程的论文中，玻恩第一次提出波函数的概率解释，他指出：

> [爱因斯坦说]波的提出只是为微粒光量子指示通道，他在一个"鬼场"(ghost field)的意义上谈论它。这决定了一个光量子(能量和动量的载体)通过某一路径的概率；然而，场本身并没有能量和动量。
>
> (1926)

这里，光波失去了它的物质性，变成了一种概率波。

[149] 我们能用几项重要的发展来填补爱因斯坦1909年的观点和玻恩1926年的观点之间的缺口。众所周知，玻尔、克拉默斯和斯莱特在他们1924年的论文中提出，电磁波的场只决定一个原子在相关空间中通过量子来吸收或发射光能的概率。这个观点实际上属于斯莱特。[2] 在这篇论文提交之前的几周，斯莱特给《自然》杂志写了一封信，在这封信中他提出，部分场将"引导离散的量子"，并且导致"[原子]以某种概率获得或失去能量，这与爱因斯坦已提出的意见非常相似"(Slater, 1924)。

在爱因斯坦1925年关于气体理论的论文中，我们可以找到爱因斯坦的波动观点与玻恩的波动观点之间的另一条联系：爱因斯坦把德布罗意的波与他自己的运动气体分子相联系，利用粒子物理学的数学形式体系，指出他的理论的基本本体是粒子而不是波，尽管在所探讨的现象中有波的一面，就像在分子神秘的相互作用中显示出波的一面那样。接下来，约当1925年的论文尽管为波粒二象性的精神所左右，但也完全是按照粒子来分析的。

关于波函数的解释，玻恩在他 1926 年的论文中打算要做的只是延伸上面提到的“爱因斯坦的观点”，把爱因斯坦关于光波与光量子之间的关系，扩展到也被他看作“鬼场”的波函数 Ψ 与可称重粒子之间的关系。玻恩的主要论据基于这两组关系之间的“完全类推”，可以总结如下：

(1) Ψ 波，作为鬼场，本身没有能量和动量，因此没有通常的物理意义；

(2) Ψ 波，与上面提到的德布罗意的观点和薛定谔的观点形成对照，完全不表征物质的运动；它们只决定物质的可能运动，或关于运动物质的观测结果的概率；

(3) 玻恩的概率意味着“在一组完全相同的、没有耦合的原子中，一种态出现的确定频率”；

(4) 因此 Ψ 波概念预设了大量独立粒子的存在。

因此，显然，玻恩把粒子看作波动力学或量子力学的基本本体。雅默把玻恩关于 Ψ 波的概率解释总结如下：

[150] $|\Psi|^2 d\tau$ 测量在单位体积 $d\tau$ 内找到粒子的概率密度，粒子在经典的意义上被设想为一个在每个时刻拥有确定位置和确定动量的点质量。

(Jammer, 1974)

玻恩是否在经典的意义上考虑粒子是一个有争议的问题。

玻恩的解释导致狄拉克的变换理论(1926b)，后者反过来导致海

森伯的不确定关系(1927)；据此，玻恩很快意识到：

> 不但必须放弃经典物理学，而且必须放弃关于实在的朴素概念，即把原子物理学的粒子看作好像是极其小的沙粒，在每一时刻有一个确定的位置和速度。
>
> (Born，1956)

玻恩承认在原子系统内“情形并非如此”。他也赞同薛定谔的观点，关于量子统计，“粒子将不被看成个体”(同上)。关于这种非经典的、非个体的粒子，玻恩解释说，之所以“保留粒子的观念”，是因为“它们表征可观测的不变量”(同上)。但出于相同的原因，波的观念也应该保留。玻恩甚至更进一步，把“概率波，甚至是 $3n$ 维空间中的概率波，看作真实的东西，当然不只是作为数学计算的工具”(1949)。他问道：“如果我们凭借这个概念不能指称某种真实的且客观的事物，我们怎么能够信赖概率预测呢？”(同上)

所有这一切都表明，雅默令人误解的说明，不但过分简化了玻恩关于量子力学的本体论立场，而且忽略了这样的重要事实：尽管玻恩的解释把一种粒子看作量子物理学的本体，但是，玻恩的解释与经典粒子本体论是不相容的，而与波动本体论共有某些特征。

对玻恩 1926 年的论文的另一个误导性说明，能在海森伯的著名论文《量子理论解释的发展》(*The development of the interpretation of the quantum theory*)中找到。在这篇论文中，海森伯声称：

> 与玻尔、克拉默斯和斯莱特的假设相比较，[玻恩的]假设包含两个重要的新特征。第一个是断言在考虑“概率波”时，我们关心

的是在一个抽象的位形空间中的过程，而不是在通常的三维空间中的过程；第二个是承认概率波与单个过程相关。

(Heisenberg, 1955)

[151] 显然，第二点与玻恩的原始意图不相吻合。正如我们上面所评述的，玻恩是在概率解释的意义上使用概率概念的。这意味着概率波包含关于一组系统的统计陈述，而不只是关于一个系统的统计陈述。关于第一个特征，海森伯的评论流于表面，且忽略了要点。事实上，玻恩提出概率解释的主要目的，准确地说，是要把薛定谔引入的奇怪的位形波恢复到通常的三维空间。玻恩本人在他1926年论文的“结语”中清楚地指出了这一点：“它[概率解释]允许保留关于空间和时间的传统观念，在这个空间和时间中，事件以一种完全正常的方式发生。”

然而，这一成就的取得不是没有代价的。正如海森伯在他1927年的论文中所清楚地意识到的，它把量子力学的解释引向如何从量子力学的表述抽取实验信息的问题。以这种方式，人们能容易地解释“波包坍缩”，但不能解释双缝实验中电子的衍射。这个实验能以一次只有一个电子穿过孔径这样一种方式进行。在这种情形中，与每个电子相联系的 Ψ 波自己与自己干涉。因此，Ψ 波一定是某种物理上真实的波，而不仅仅是我们关于粒子知识的一种描述。

作为对玻恩解释所面临的这种困难的一个回应，海森伯把概率波重新解释为与上面提到的单个过程相关联的波，并把概率波看作一种新的物理实在，与一种可能性或“潜在”(*potentia*)相联系，是“实在的某种中间层次，位于物质的整体实在与观念的智力实在之间的中间状态”(Heisenberg, 1961)。按照这种观点，一个由某个波函数

(或这种函数的一个统计混合)描述的封闭系统，在性质上是潜在的而不是现实的，因此，在测量过程中发生的“波包坍缩”是一个从潜在到现实过渡的结果，这种过渡由封闭系统与外部世界(测量仪器)的关联所产生(Heisenberg，1955)。

“潜在”这一形而上学概念的引入，标志着基于粒子本体的概率解释所面临的困境，是对波函数的实在论解释或波动本体的一个重要让步。事实上，海森伯早在1928年就作出了这一让步；其时约当、奥斯卡·克莱因和维格纳证明，一个 n 粒子系统的薛定谔波描述与遵循量子统计(玻色—爱因斯坦统计和费米—狄拉克统计)的粒子的二次量子化描述等价。[3] 例如，海森伯其时认为粒子绘景和波动绘景只是同一个物理实在的两个不同方面(参阅 Jammer，1974)。他也意识到： [152]

> 只有位形空间中的波，即变换矩阵，是惯常解释中的概率波，而三维物质波或辐射波则不是。根据玻尔、约当和维格纳，后者有与粒子刚好一样多(和刚好一样少)的“客观实在性”；它们与概率波没有直接的关联，但像麦克斯韦场那样有连续的能量密度和动量密度。
>
> (Heisenberg，1955)

这也就是说，海森伯接受存在着不同于概率波的物质波，它们有与粒子一样的本体论地位。

这里不太清楚的是“物质波”指什么。如果它们可以解释为可称重物质的某种实体波，那么就是对波函数的实在论解释的一大让步，即使“三维的”限制似乎暗示只有单个粒子的德布罗意波能被看

作“客观实在”。另一方面，如果“物质波”只是“辐射波”的同义词，那么问题是：何时把不同于概率波的辐射波(如光子的德布罗意波)看作“客观实在”？对于只有一个电子或其他可称重粒子的三维德布罗意波，情况又将如何？它们仍然是概率波，或是客观实在吗？事实上，海森伯和所有在量子场论领域工作的其他物理学家不断地在这两种意义之间来回变换。使用意义不明确的“物质波”概念，实际上是所有关于量子场论本体概念混淆的起源。但是，另一方面，确切地说正是这种含糊性，使得量子场论的发展有可能建立在物质波场的概念基础之上。这一概念任何清楚的、无歧义的使用将造成严重的概念困难。因此，这种含糊性似乎反映和标志着人们在概念上理解量子场论本体时所面临的最为深刻的困难。这也解释了为什么这种含糊性已经奇怪地持续了60多年，而没有显示出任何被阐明的迹象。

7.2 量子化

关于量子场论本体的讨论，在很大程度上，也依赖于人们对于量子化(quantization)概念的理解：量子化意味着什么？量子化的数学程序意味着什么？量子化的物理含义与数学程序之间的关系又是什么？可以说，文献中盛行的一些影响深远的混淆，都可追溯到对量子化这一概念的误解。量子化的物理思想虽然早在1900年随着普朗克的黑体辐射理论的提出就已出现，但是量子化的数学程序只是在1925年海森伯、玻恩、约当，特别是还有狄拉克发展的量子力学出现以后才被建立起来。最初，数学程序只是用作实现量子化思想的物理含义的一个工具。但是，一旦数学程序建立起来以后，它就呈现为一种独立的存在。

[153]

因此，易犯的错误就是忘记、忽略或割断量子化的物理方面和数学方面的密切联系，把形式化的程序看作量子化的本质，看作一个模型，在其中或基于其上，能实行量子化、二次量子化和场的量子化的讨论，因此给出了使人误入歧途的解释。在这节我将通过简明的历史考察，来确定量子化概念的本义，澄清其主要误解。

对普朗克而言，量子化为热辐射定律所要求，只意味着一个简单振子的能量的量子化：在一个热原子中进行简谐振动的电子的统计行为中，只有那些能量是量子的整数倍的振动态才应该被考虑。这里，能量的量子化是原子中电子的一个属性，由与辐射相互作用的振子来表示。

对玻尔来说，量子化意味着原子中电子的周期轨道的量子化，量子化的电子轨道与离散的定态相联系。因此，与定态相联系的力学运动将由量子化条件(QC)从这种运动的连续流形中选择出来。根据玻尔，量子化条件应该考虑为“对普朗克有关谐振子可能能量值的最初结果的一种合理推广”，因为这里，力学运动的量子化不但涉及能量的量子化，而且涉及原子系统中动量和角动量的量子化。

从数学上讲，玻尔的量子化条件适合于运动的每个周期坐标 Q，能用 $J=\oint P\mathrm{d}Q=nh$ 表示，其中 P 是与 Q 相对应的正则共轭动量(如果 Q 是一个角，那么 P 可以是角动量)和一个周期的积分。选择这个积分 J 的最令人信服的理论依据由埃伦费斯特(1916)给出，埃伦费斯特表明，如果系统经受一个慢的外部微扰，那么 J 不变，因此非常好地与一个非连续“跃迁”量 nh 相匹配。

玻尔的量子化观念，既在它的物理学方面(作为原子中电子的力学运动的量子化)也在它的数学方面($J=\oint P\mathrm{d}Q=nh$)，构成了量子 [154]

化的两个学派成长的起点。第一个是德布罗意—薛定谔学派。根据德布罗意(1923a, b)，运用表达式($P = h/\lambda$, $J = \oint P\mathrm{d}Q = 2\pi rh/\lambda = nh$, $2\pi r/\lambda = n$)，量子化条件可以解释为：与一个运动电子相联系的波，其围绕轨道周线的波长数必定是一个整数。对薛定谔来说，正如其标题为“作为本征值问题的量子化”的四篇开创性的论文所表明的，量子化能约化为本征值问题(1926b, c, e, f)。对于波函数 Ψ (与原子中某些振动过程相联系)的任意变量来说，如果(占据所有坐标空间的)哈密顿函数的积分是固定的，那么 Ψ 函数能被选作适合于一组离散的能量本征值，因此，被选的 Ψ_s 本身变成与本征值相对应的本征函数。

从数学上讲，薛定谔的变分问题要比玻尔—索末菲的量子化条件 $\oint P\mathrm{d}Q = nh$ 复杂得多。然而，其指导原则与后者存在的亲缘关系，能从“定态积分”的观念与埃伦费斯特绝热不变假说的相似性中找到。至于与德布罗意的概念之间的关系，薛定谔本人指出，他直接从德布罗意的如下结论中获得灵感：与相波的空间分布相关联，总存在着一个沿着电子轨道测量到的整数。薛定谔继续指出，他与德布罗意的“主要不同是德布罗意思考前进波(progressive wave)，而我们被引导到稳定的本征振动”(1926b)。

第二个学派以海森伯(1925)，玻恩和约当(1925)，以及狄拉克(1925)为代表。在这个学派中，量子化的基本物理含义与玻尔概念中的物理含义完全相同：原子中的力学运动被量子化。但是，得到这个量子化的数学表达式的指导原则超越了埃伦费斯特的绝热假说。这决定性的一步由海森伯(1925)完成。海森伯在唤起人们注意整个相空间积分(它表示原子中的力学运动)的量子化之余，还唤起人们注

意原子物理学中动力学量(或物理可观测量)的量子本质。这些量依赖于两个量子化定态的跃迁量，因此应该用矩阵来表示，玻恩和约当(1925)两个月后证明了这一点。玻恩和约当表明，考虑这个可观测量，当引入 P 和 Q 的傅里叶展开($P = \sum_{-\infty}^{\infty} P_\tau e^{2\pi i\nu\tau t}$，$Q = \sum_{-\infty}^{\infty} Q_\tau e^{2\pi i\nu\tau t}$)时，玻尔—索末菲量子化条件 $\oint P\mathrm{d}Q = \oint_0^{1/\nu} PQ\mathrm{d}t = nh$ 能转化为 $1 = 2\pi i\sum_{-\infty}^{\infty}(\tau\partial/\partial J)(Q_\tau P_{-\tau})$。考虑克拉默斯的色散关系(1924)，这一色散关系宣称总和 $\sum_{-\infty}^{\infty}(\tau\partial/\partial J)(Q_\tau P_{-\tau})$ 必定对应于 $1/h\sum_{-\infty}^{\infty}[Q(n+\tau, n)P(n, n+\tau) - Q(n, n-\tau)P(n-\tau, n)]$，于是，量子化条件就变为 $\sum_{-\infty}^{\infty}[P(nk)Q(kn) - Q(nk)P(kn)] = h/2\pi i$，或者用矩阵表示为 [155]

$$\boldsymbol{PQ} - \boldsymbol{QP} = (h/2\pi i)\boldsymbol{1} \quad (1)$$

1925 年，狄拉克通过重新解释经典哈密顿动力学方程，获得了 q 数的相同的非对易关系(1925)。

动力学量的非对易关系的数学表达式的建立，标志着原子中力学运动的量子化的数学描述的完成。在这一数学描述基础之上，提出了一种量子化程序：用相应的 q 数代替经典哈密顿表述中的正则共轭动力学量，并使它们受制于非对易关系(1)，这样就获得了量子化表述。

数学程序[原本只是]作为代用品和附属物，[现在反倒]迅速变成量子化模型，在此基础上，人们给出了各种量子化解释。因此，这深刻地影响了量子物理学随后的发展，不但在其数学描述上，而且在其概念的物理解释和本体论解释上。然而，这里我想强调的是，我们应该记住这种数学程序的基础、前提、真正含义和限制。

海森伯的决定性起点是玻尔将辐射频率(跃迁量或所谓的可观测量)与两个定态联系起来的频率条件。因此，海森伯“可观测量”的观念预设了原子中量子化定态的存在。最初，海森伯，玻恩和约当，还有狄拉克，他们关于量子化条件的表达只是玻尔的量子化条件的重写。因此，如果不把它与原子中的量子化定态相联系，那么这一量子化条件的表达对于考虑量子化程序是没有意义的。

然而，这一限制不久就被薛定谔和狄拉克随后开创的发展去除了。薛定谔在回顾波动力学和矩阵力学之间的关系时认为，矩阵的本质特征是它表征了一个作用于矢量(列矩阵)的线性微分算符。用这种方式，薛定谔(1926e)走向他的算符微积分。如果坐标 q 作为一个普通变量，与在它自己的表象(Q 表象)中表示的一个动力学变量

[156] (比如说 Q)的对角矩阵相对应，那么相应的正则共轭动量 p 将随处被算符 $h/\mathrm{i}(\partial/\partial q)$ 取代。以这种方式，量子化条件(1)变成了一个平凡的算符恒等式。于是薛定谔的量子力学思想与此类似。系统一个确定的态由一个在某种 q 表象中的波函数 $\Psi(q)$ 定义，q 是系统的一个可观测量。动力学变量(或可观测量) A 能用一个线性算符表示，并且使 A 作用于 $\Psi(q)$ 上，$A\Psi(q)$ 意指表示系统中一个新态的新波函数。于是，如果这个新函数除了一个常数因子外，与 $\Psi(q)$ 完全相同，即 $A\Psi(q) = a\Psi(q)$，那么 $\Psi(q)$ 就称为 A 的一个本征函数，常数 a 是一个本征值。本征值的完全集表征了该算符 A，表示了这个可观测量的可能数值，它们可以是连续的，也可以是离散的。本征值全部是实数的算符称为厄米算符。显然所有物理量都必须用厄米算符描述，因为本征值被假定表示对一个物理量进行测量的可能结果。

在薛定谔这一思想的基础之上，狄拉克(1927a)获得了对量子力

学各种可能的表象的洞见。他承认最初矩阵力学采用的形式是能量表象，最初波动力学采用的形式是坐标表象。这只是可能的表象中两种特殊的情形。除此以外，还有角动量表象、粒子数表象等等。狄拉克指出，在他自己的符号体系中，如果动力学变量 G 最初采用 α 表象 $\langle\alpha'|G|\alpha\rangle$，那么它的 β 表象能通过幺正变换获得：$\langle\beta'|G|\beta\rangle = \iint \langle\beta'|\alpha\rangle \mathrm{d}\alpha \langle\alpha|G|\alpha'\rangle \mathrm{d}\alpha' \langle\alpha'|\beta\rangle$，其中变换函数 $\langle\alpha'|\beta\rangle$ 确切地说是表征这一动力学变量的算符的本征函数。例如，薛定谔波动方程(作为一个能量本征方程)的本征函数恰好是从坐标表象到能量表象的变换函数。这就是所谓的一般变换理论，因为 α 和 β 可以是对易的可观测量的任何集合。

量子化条件(1)的基础因此扩宽了。首先，所考虑的物理系统不再局限于具有离散能级的原子，因为薛定谔波动方程使得处理连续的本征值成为可能。自由粒子和碰撞问题因此也进入到量子力学的范围。第二，狄拉克变换理论使得可由量子力学处理的系统的物理态，不再局限于系统的用能量本征函数表示的能态(例如原子的量子化定态)，并且各种用表征能量之外可观测量的算符的本征函数来表示的物理态，也都可以处理。

考虑到这些推广，量子化条件(1)似乎已成为量子化的普适条件。[157] 然而，仍有一个苛刻的限制保留在所有这些推广之上：只有那些真正可观测量的算符对、矩阵对或动力学变量对，受量子化条件(1)支配；通过这一量子化条件，可以获得一个系统的力学运动的真正量子化。“真正的可观测量”，我意指是一个用厄米矩阵表示的可观测量，能找到一个表象并使矩阵在该表象中是对角化的。

可能有人把这个限制看作是不切实际的，因为量子物理学中的动

力学变量似乎等同于可观测量。但实际上这在量子场论情形中，甚至在量子力学的多体问题的情形中，却是不真实的。奇怪的是，我们会发现量子场论的所有先驱者和奠基者，诸如约当(1925b，c)、狄拉克(1927b，c)、奥斯卡·克莱因(和约当，1927)、维格纳(和约当，1928)、海森伯和泡利(1929)，以及费米(1932)，都是让一组不可观测量(诸如粒子数表象中的产生与湮灭算符，见7.3节)服从量子化条件(1)的各种变种，然后声称得到了一个量子化理论或二次量子化理论。这个给出量子场论的物理学解释的量子化程序，因此成为引起量子场论本体论混淆的另一个根源，这一点我们将在下一节中详细说明。

量子化概念的另一个发展与量子场论关系更为密切。这里考虑的物理系统是连续场而不是离散粒子。这一发展的开创者是爱因斯坦。在普朗克物质振子的能量的量子化基础之上，爱因斯坦提出"由局域在空间不同点的有限数目的能量子组成的光的能量，只能作为一个单元而被产生或吸收"(1905a)。这个思想为埃伦费斯特所接受，埃伦费斯特在其论文中提到，由一个简谐振子表示、处于频率 ν 的简正模式中的电磁场能量，只能是 $h\nu$ 的整数倍(1906)。埃伦费斯特的思想，即辐射本身可以看作简单振子系统，其中每一个简单振子表示一个平面波的振幅，为德拜(1910b)、约当(1925)和狄拉克(1927b，c)所继承。因此场的能量的量子化可以通过把普朗克假说应用到场振子而轻易获得(德拜)，或者通过应用量子力学中用于处理物质振子的数学方法而轻易获得(约当和狄拉克)。

然而，爱因斯坦关于场的量子化思想实际上要比这更为深刻。在上面提到的同一篇论文(1905a)中，爱因斯坦推测了辐射本身是由"独立的能量子组成"的可能性。爱因斯坦的光量子假说进一步被他自己

的贡献(1909a，b，1916c，d，1917b)和康普顿的实验(1923a，b)所支持，这在第6章中曾经评论过。在此，我只想给出两点评论： [158]

(i) 爱因斯坦的光量子思想涉及一种新的量子化，是物质场的量子化而不是粒子或场的力学运动的量子化。

(ii) 对于场的量子化来说，不存在与爱因斯坦的光量子假说相伴随的数学程序。在场的量子化中，用作数学程序的只是在力学运动的量子化中所用的类似程序。因此，有一个概念上的缺口需要填补，并且需要给出这一严格类推的正当理由。

现在，让我们转向关于这些问题的讨论。

7.3 场的量子化

正如我们上面提及的，场的量子化概念有两种意义：(i)场能(或更一般地说，场的力学运动)的量子化，这种量子化与量子力学中处理的粒子的力学运动的量子化相类似；(ii)场作为一个物质实体的量子化。

如果我们接受麦克斯韦的思想，认为能量一定存在于物质之中，同时也接受爱因斯坦的思想，认为作为电磁场能量载体的物质不是机械以太，而是电磁场本身；电磁场作为一个独立存在的实体，从本体论上看等同于可称重物质，那么“场”不能被看作“场能”的同义词，但前者与后者相关，如同物主之于他的财产。用这种观点，由意义(i)推不出意义(ii)，反之亦然。这个关于推不出的断言为如下事实所支持：在连续或非连续的能量与其载体之间没有必然的联系，这一事实就粒子的情形来说很容易检验。一个电子，当它自由的时

候，能量是连续的；当它在原子中时，能量就是离散的。

然而，在量子场论的历史中，在“场的量子化”的含糊标题下，这两种意义经常被简单地混为一谈；人们从这样一个含义不明确的题目中，粗心地推出一些没有根据的推论。由于场的量子化概念是量子场论概念基础的基石，因此仔细地对这一概念的形成进行历史研究看来是十分必要的，下面就是对此所做的一个尝试。

通常认为，场量子化的思想可以追溯到埃伦费斯特的论文(1906)和德拜的论文(1910b)。[4] 根据这两位作者，如果我们把空虚空间中[159]的电磁场看作一个谐振子系统，把普朗克关于物质振子能量的量子假说应用于场振子，那么不需要用物质振子，就能推出普朗克能谱。验证这些启发性思想的恰好就是普朗克思想本身，即被观测到的黑体辐射谱。显然，这里场的量子化只意味着场能的量子化，即只是意义(i)。

约当的工作在量子场论的历史上占据着显著的位置。在与玻恩合作的一篇论文(1925)中，约当不但给出了量子化条件(1) $\boldsymbol{PQ}-\boldsymbol{QP}=(h/2\pi i)\mathbf{1}$ 的证明，而且给出了与现在的讨论更为贴近的原创性思想，即电场和磁场应该被看作动力学变量，由矩阵表示，并服从量子化条件，以使辐射的偶极发射经典公式可以纳入量子理论。在与玻恩和海森伯合作的论文(1926)中，如同埃伦费斯特和德拜一样，约当把电磁场看作一组服从量子化条件的谐振子。他以这种方式推导出黑体辐射场中的均方能量涨落。[5]

由于爱因斯坦从一个波场中的涨落定律推证出了微粒光量子概念的合理性，因此上述成功使约当确信，“令人烦恼的爱因斯坦光量子问题，可以通过把量子力学应用于麦克斯韦场本身而予以解决”(Jordan，1973)。这里，应该注意两点。(i)要被量子化的场是实体

的且连续的麦克斯韦场，(ii)微粒光量子可以作为这个场的量子化的一个结果而获得。考虑到以下事实，即爱因斯坦的光量子不但拥有确定的能量 $h\nu$(正如他最初提出的)，而且拥有确定的动量 $P = hK$ (正如他在 1916 年的论文中意识到的，见 6.2 节)，因此，爱因斯坦的光量子是完全成熟的粒子，而不仅仅是能量子。这里，场的量子化显然是在意义(ii)上使用的。

为了判断约当的上述确信是否正当，我们必须阐明约当实际上做了什么。对他 1925 年的论文所做的考察表明，他做的工作只是把海森伯—玻恩量子化条件应用于波场。请记住，海森伯—玻恩量子化条件预设了：(i)原子的量子化定态的存在，(ii)可观测量与原子的两个态相关联(见 7.2 节)。因此我们发现，约当对海森伯—玻恩量子化条件的应用隐含着他心照不宣地采纳了两个假设，即将海森伯—玻恩量子化条件的预设推广至麦克斯韦场：(i)麦克斯韦场的能态也是量子化的，(ii)场变量作为动力学变量总是与场的两个态相关联。他对这一推广没有给出清楚的理由，尽管原子的能量改变和场的能量改变之间的密切联系似乎间接地支持这样一个推广。但是无论如何，用这两个假设，约当能“证明”的只是场的能态的量子化，这确实是他已预示的一个结果，但与麦克斯韦物质场本身的量子化没有关系。事实上，约当的论文的意义在于它的数学方面：借助于振子模型，他为实现场能的量子化的物理假说提供了一种数学表述。 [160]

相对而言，1925 年的情形还是比较简单的。尽管德布罗意推测物质波存在，但是情形并没有多大的改变。在 1925 年，当人们提到场时总是指物质场。当薛定谔 1926 年的工作取得进展后，情况就变得复杂多了。薛定谔的波很快就成为被量子化的基本实体。甚至电磁波被看作一种与光量子相关的薛定谔波(见下面关于狄拉克论文的

讨论)。然后，正如我们已经注意到的，关于薛定谔波函数有两种解释：一种是实在论的，另一种是概率论的。它们各自为量子场论假定了不同的基本本体。除了对波函数的解释的混淆之外，还有另一种关于场的量子化的两种意义的混淆。因此，对场的量子化的讨论需要仔细分析，这样才能澄清这一复杂的情形。这将涉及对流行的误解的批判和对包含在量子场论奠基者的原初论文中使人误入歧途的假设的澄清。

我将从狄拉克的论文“关于辐射的发射和吸收的量子理论”(1927b)入手，这篇论文由于其中包含有“二次量子化的发明”，而被看作量子场论发展的萌芽(Jost，1972)。关于这篇论文，有许多广泛传播的误解。首先，狄拉克继续进行量子化的场是什么？关于这个问题，约当在他的回忆录(1973)中告诉我们，当狄拉克的论文刚发表时，他和玻恩把场看作“粒子的本征函数”，或者“单粒子的薛定谔场”。事实上，约当一生都坚持这种理解，并把他 1925 年的论文以及随后关于量子场论的论文，也都看作是以薛定谔场作为它们的起点。与之相似，温策尔(Gregor Wentzel，1960)在他关于 1947 年前的量子场论历史的权威论文中，也声称场被看作“一个(复)薛定谔波函数”或“‘未受扰动的态’的概率幅”。

狄拉克和约当形成对照，约当在他处理辐射场的论文(Jordan，[161] 1925；Jordan and Born，1925；Jordan，Born and Heisenberg，1926)中，混淆了辐射场和薛定谔波，而狄拉克在他一篇论文的“导言和概要”中则清楚地区分了这两种场：

> 首先，光波总是实的，而与光量子相联系的德布罗意波……必定包含一个虚指数。一个更重要的不同是：它们的强度是以不同

的方式来解释的。与单色光波相联系的每单位体积的光量子数等于波的单位体积的能量除以单个光量子的能量 $h\nu$。另一方面，一个幅度为 a 的单色德布罗意波……必须被解释为表示每单位体积有 a^2 个光量子。

(1927b)

他也指出，上述多体(多光量子)薛定谔波的解释是波函数的一般概率解释的“一种特殊情形”。

据此，如果(ξ'/α')或 $\Psi_{\alpha'}(\xi'_k)$ 是一个原子系统(或一个单粒子)的态 α'的变量 ξ_k 中的本征函数，那么 $|\Psi_{\alpha'}(\xi'_k)|^2$ 是每个具有值 ξ'_k 的概率。

(同上)

现在很清楚，狄拉克的言下之意有两点。(i)他把薛定谔波等同于概率波，并把它们与粒子本体相联系。(ii)他把真实的物质波和薛定谔波区分开来，声称“没有与电子相联系的这种(物质)波”。这里不很清楚的是他对于光波(或光波中的能量子)和光量子之间关系的看法。他就这一关系的阐述对理解量子场论早期的概念情形至关重要，因此我们必须深入其细节。

在这篇论文中狄拉克潜在的观点之一是，“在波和相互作用的光量子描述之间有一种完全的和谐”。为了说明这一点，他实际上首先根据光量子的观点建立起理论，然后表明粒子表述能自然地转换为一种波的形式。

狄拉克借助于所谓的“二次量子化”(second quantization)程序实施了第一步。他从一个有 N 个相似的独立系统的微扰系综的本征函数 Ψ 出发，并用未受微扰系综的本征函数 Ψ_r 来表示 Ψ：$\Psi=\sum_r a_r\Psi_r$，然后获得系综处于态 r 的可能数 N_r：$N_r=|a_r|^2$。通过引入变量 $b_r=a_r\exp(-\mathrm{i}\omega_r t/h)=\exp(-\mathrm{i}\theta_r/h)N_r^{1/2}$ 和 $b_r^*=a_r^*\exp(\mathrm{i}\omega_r t/h)=\exp(\mathrm{i}\theta_r/h)(N_r+1)^{1/2}$，并且令 b_r 和 $\mathrm{i}hb_r^*$ 是满足量子化条件的正则共轭“q 数”：$[b_r,\mathrm{i}hb_s^*]=\mathrm{i}h\delta_{rs}$，

[162] 他获得了一个理论；在这个理论中，相互作用哈密顿量 F 能用 N_r(粒子数)写为[6]：

$$F=\sum_r W_r N_r+\sum_{rs}V_{rs}N_r^{1/2}(N_s+1-\delta_{rs})^{1/2}\exp[\mathrm{i}(\theta_r-\theta_s)/h] \tag{2}$$

这里，起点是多粒子薛定谔波，基本本体是光量子——一种携带能量和动量的稳定粒子。关于狄拉克的光量子观点，原文中这段话表述得最为清楚：

> 光量子有一种特性，当它处于一个定态，即动量以及能量为零的零态时，它显然不再存在。当一个光量子被吸收时，可以被认为是跃迁到了零态；当它被发射时，可以被认为是从零态跃迁到了一个它的物理性质明显的态，因此表现为产生。
>
> (1927b)

第一步中所做的一切与场的量子化无关，因为，首先，根本就没有任何真实的场(在 7.1 节中提到，多粒子薛定谔波函数的任何实在论解释均面临严重的困难)，其次，还不存在波场的量子化或二次量

子化的恰当定义。

这个结论性陈述的推理很简单，但经常为量子物理学家和量子物理学史学家所忽略。请注意，量子化条件应该遵循的一个严格限制是：只当它涉及的动力学变量是由厄米算符表示的可观测量时，量子化条件才有意义。因此狄拉克的假设——人们能把表征概率场幅度的非厄米算符 b 和 $\mathrm{i}hb^*$ 看作是正则 q 数，且服从量子化条件：$[b_r, \mathrm{i}hb_s^*] = \mathrm{i}h\delta_{rs}$ 或 $[b_r, b_s^*] = \delta_{rs}$ ——显然是误导。事实上，“二次量子化”这一误导的称谓，只不过是对粒子力学运动的量子化条件的等价表述，它是借助于在粒子数表象上很方便的“产生”和“湮灭”算符 $a^*(b^*)$ 和 $a(b)$ 而实现的。特别是，“产生”和“湮灭”算符所满足的对易关系与概率场的量子化没有关系，因为没有涉及普朗克常量（$[b_r, b_s^*] = \delta_{rs}$）。它们只是这些算符的代数性质。

现在让我转到狄拉克的第二步。只有在这一步他才处理了场的量子化。这一步的起点是经典辐射场。狄拉克把辐射场分解为傅里叶分量，考虑每个分量的能量 E_r 和相位 θ_r，以形成一对描述场的正则共轭动力学变量，并服从标准的量子化条件 $[\theta_r, E_s] = \mathrm{i}h\delta_{rs}$。[7] 这里，辐射场真的被量子化了。因此，问题在于：在什么意义上，辐射场被一次或二次量子化？ [163]

这个答案近在咫尺，如果我们注意到以下两点。首先，场的量子化受场振子量子化的影响（因为场的每个谐波分量实际上是一个简谐振子）；其次，场振子的量子化，作为振子的一种量子化，只能按照第一种意义来理解，即场能或场振子的量子化。

在这个解释中，只有当波场的能量子（这里，基本本体是场本体，能量子只是描述场的性质的一种方式）与由薛定谔概率波描述的光量子（这里，基本本体是粒子本体，根据玻恩，薛定谔波只是归因

于有某些性质的粒子出现的概率)同一时，狄拉克关于假定场变量 E_r 和θ_r(是 q 数)“直接把光量子的性质给予辐射”的断言才能被证明是正确的。狄拉克通过把多粒子问题的薛定谔方案中每个定态的 N_r' 个光量子数，看作是他 1927 年论文(1927b)第 263 页上关于场分量中能量的 N_r 个量子数，心照不宣地假定了这种同一性。在那篇论文中，狄拉克根据他的新理论重新推导了爱因斯坦关于辐射和吸收的概率系数。用这种方式，他获得了一个描述原子和电磁波的相互作用的哈密顿量。

通过让人们注意到这个哈密顿量与原子和一组光量子相互作用的哈密顿量相等同，狄拉克声称“波动观点因此与光量子观点一致”。这个等同性的主张给人印象如此深刻，以至于许多物理学家，包括狄拉克本人，把它看作“辐射的波粒二象性之谜已经解决”的一个证明(Jost，1927；见 Dirac，1983)。

然而，这里我们应该注意：

(i) 能量子作为波动系统中的峰是可以创生和破坏的。因此，作为一种副现象，它们预设了作为基本本体的场的存在，尽管它的数目能被用作确定场振子的态和场的位形的参数。

(ii) 在狄拉克的头脑中，光量子是一种独立的、永久的粒子，携带着能量、动量和极化，并处于由这些参数确定的某个态。它们的产生或湮灭只是它们跃迁出或跃迁进零态的表现。这里粒子是基本的物质，概率波只是一种便于计算的智力设计。

[164]

(iii) 狄拉克将能量子与光量子视为同一，因此是一个重大的概念跳跃，通过这一概念跳跃，他把场的性质(场的量子化的能量)与一种离散的物质(光量子)的存在相等同，并把连续的

物质(场)与非真实的概率波混而为一。然而却没有给出这种概念跳跃的正当理由。

如果我们认真地理解这些评论，那么狄拉克关于场的量子化能够实现从场描述到光量子描述的转换的断言，在它的本体论方面就不能认为是可靠的。原因在于这一转换严重依赖于那种等同性，而这种等同性是一个与转换本身一样根本和重要的假设。

狄拉克在他接下来的一篇论文(1927c)中表明，在他的方案中，就量子化而言，与广泛流行的观点相反，他认为被量子化的场总是指经典辐射场，而不是薛定谔波场，并且也与广泛流行的观点相反，他认为量子化仅仅指经典场的量子化，而不是薛定谔波的二次量子化。因而问题在于：为什么这一误导人的观点如此流行，以至于它甚至为量子场论的大多数奠基者诸如约当、奥斯卡・克莱因、维格纳、海森伯、泡利和费米所接受。对他们来说，这个误导性的观点是他们基本论文中的根本思想和起点。因此，这个观点已经深刻地影响了量子场论的概念发展。原因有两个方面：第一，它源于狄拉克原文中对场量子化的特殊处理。第二，它源于在与满足费米—狄拉克统计的粒子相联系的费米场论中所包含的一个更为深刻的概念混淆。

让我们首先来考察第一个原因。在狄拉克获得辐射场的量子理论的论文中(1927b，c)，主要步骤如下：(A_1)从经典辐射场出发；(A_2)把场看作动力学系统，即把能量和场的每个分量的相位 E_r 和 θ_r'(或等效地说，是能量子数 N_r 和正则共轭相位 $\theta_r = h\nu\theta_r'$)，看作描述场的动力学变量；(A_3)假定 N_r 和 θ_r 是满足量子化条件的 q 数，$[\theta_r, N_s] = ih\delta_{rs}$。然而，在他的原始论文中，特别是在1927年的一篇论文(1927b)中，狄拉克把这一条推理线路与另一条也由三个主要

步骤所组成的推理线路混淆了起来。另一条线路的三个主要步骤是：(B1)从多光量子系统的薛定谔波 $\Psi=\sum_r b_r\Psi_r$ 出发；(B2)把概率幅 b_r 和它的共轭量 ihb_r^* 看作满足量子化条件的正则 q 数，$[b_r, ihb_s^*]=ih\delta_{rs}$；(B3)用另一对满足量子化条件 $[\theta_r, N_s]=ih\delta_{rs}$ 的正则 q 数，N_r(光量子数)和 θ_r(薛定谔波的相位)，通过以下公式把 b_r 和 b_r^* 联系起来，

[165]

$$\begin{aligned} b_r &= (N_r+1)^{1/2}\exp(-i\theta_r/h)=\exp(-i\theta_r/h)N_r^{1/2}, \\ b_r^* &= N_r^{1/2}\exp(i\theta_r/h)=\exp(i\theta_r/h)(N_r+1)^{1/2} \end{aligned} \quad (3)$$

导致这种混淆的原因似乎是双重的。首先，在本体论上，狄拉克混淆了场本体论中的能量子和粒子本体论中的光量子(1927b, p.263；1927c, p.715)。当他把电磁场的每一分量看作"是与某一种类型的光量子相联系"时，他甚至表露出一些把真实的电磁场解释为光量子的德布罗意波的倾向(1927c, p.714)。这种电磁场的粒子本体论解释，当然与 1909 年爱因斯坦的观点、斯莱特的观点、玻恩和约当的观点相一致(见 6.2 节)，并且也为后来量子场论的粒子本体论倾向的解释铺平了道路，现在这种解释在物理学家中间占主导地位。其次，从技术上讲，如果不依赖于光量子绘景并把能量子数等同于光量子数，狄拉克就无法在矢势和每个场分量中能量子数之间获得一个定量关系。

这种混淆给予人们一种错误的印象。约斯特(Jost)的观点很有代表性，他把狄拉克的量子场论纲领概括为："让矢势量子化并取代经典相互作用。"(1972)

矢势的量子化思想与"二次量子化"的思想联系密切，因为在两种情形中矢势的振幅 a_r 或概率幅 b_r 被假定是一个 q 数。b_r 是一个

q 数的假设起源于狄拉克的论文(1927b, c)。然而，狄拉克没有明确假定 a_r 也是一个 q 数，尽管 a_r 在他的量子场论方法中所起的作用，与 b_r 在他 1927 年关于多粒子方法的论文(1927b, c)中所起的作用完全相同。然后，从 1929 年起，费米在他极具影响力的演讲和论文(1929, 1930, 特别是 1932)中，把狄拉克关于 b_rs 的思想扩展到了 a_rs，从而使得这一思想对于量子物理学共同体来说变得可以接受。

然而，从概念上讲，所有这些思想都是混乱的。首先，振幅 a_r 或 b_r 不能被看作在它原初意义上的 q 数，因为它不是一个可观测量，并且不能用一个厄米算符表示。事实上，$a^*(b^*)$ 或 $a(b)$ 能被解释为振子模型中的产生或湮灭算符，在振子模型中，它通过一个量子产生或退激一个振子的激发。因此振幅没有真实的本征态和本征值。也就是说，它们不能用厄米算符表示。其次，由于 $[b_r, b_s^*] = \delta_{rs}$，因此 b_r(或 a_r)不是一个 c 数。但这与量子化没有关系。[166]

为了清楚地说明这一点，我将强调下面的事实。在辐射的量子理论中，光量子数 N_r 和波场相位 θ_r 作为可观测量一定是 q 数。如果情况是这样，那么从变换方程(3)可得出 b_r 不是一个 c 数。但是，反之也不为真。N_r 和 θ_r 是 q 数不会自动地从非对易关系 $[b_r, b_s^*] = \delta_{rs}$ 中得出。这一陈述为下面的观测结论证明是正确的：方程(3)的两条线路中的第二个方程不能从 b_r 和 b_r^* 不是 c 数的事实获得，而必须从 N_r 和 θ_r 是正则 q 数(Dirac, 1927b)这样一个特殊的假设推出。简言之，b_r 和 b_r^*(或 a_r 和 a_r^*)不是 c 数的事实，承袭于 N_r 和 θ_r 是 q 数的事实。这是这个理论的量子本性的一个必然结果，而不是一个充分条件。仅仅通过把振幅看作一个非 c 数，经典辐射理论就不能被量子化。另一方面，在量子理论中，振幅总是可以为一个不是 c 数的量所表示。

因此，出于两个原因，“二次量子化”的思想是不适当的。原因之一：在量子理论中，它不是一个额外的量子化，而是一个形式变换。原因之二：它不能用作经典理论量子化的一种方法。

现在，让我们转向另一个原因：为什么狄拉克量子场论的起点是薛定谔波这个误导人的观点会如此流行。正如我们上文所谈到的，和量子场论的所有其他奠基者一样，狄拉克把薛定谔波解释为一种概率波。在本体论上，薛定谔波函数的概率解释与场论不一致，因为前者的基本本体是离散粒子，而后者的基本本体是连续场。尽管有这一明显严重的困难，物理学家仍然相信量子场论的起点一定是薛定谔波函数，它必须被看作一种经典场，服从“二次量子化”。事实上，“量子—经典—(二次)量子”这样一个概念上的诡计，已成为量子场论被接受的概念基础。我认为，这个中原因当然要比关于物理学家包括狄拉克本人的混乱深刻得多，并且与费米子场的特性密切相关。由于这个原因，让我们更为切近地查看一下这个令人迷惑的观点。

狄拉克在其辐射理论中处理的是一个多粒子问题。根据新出现[167]的波动力学，这个问题由高维位形空间中的薛定谔波函数所描述。从这样一个位形波出发，狄拉克做了一个表象变换，并获得了一个粒子数表象；在粒子数表象中，除了通常的三维空间外，不需要任何位形空间。这种被不恰当地称为“二次量子化”的方法，完全基于粒子本体论，与场论没有任何关系。对于多粒子问题，狄拉克还发展了另一种方法，即量子场论。他的量子场论显然相似于二次量子化的方法，但在本体论上却不相同，因为它的起点是通常三维空间中的经典场，而不是位形空间中的薛定谔波。诚然，狄拉克的量子场论也预示了离散粒子的存在。但这里粒子被认为是连续场的量子。也就是说，量子场论中粒子的地位被降低为副现象：它们只是基底物质

(即场)的表现形式，能创生和破坏，因此与永久存在的场完全不同。

对于物理学家来说，把位形波看作某种真实的东西是奇怪的和困难的。大多数物理学家憎恶它，企图只是把它用作一种计算手段，而不是物理实在的一种表象。玻恩的概率解释和狄拉克的“二次量子化”可以看作是试图让日常空间避开位形空间的范例。狄拉克的量子场论提供了着手研究多粒子问题而不涉及位形空间的另一个例子，这是那一时期的中心任务之一。

由约当和维格纳(1928)跨出的下一步是，将狄拉克的思想从玻色子(光量子)情形推广到费米子情形中去。对约当来说，量子场论方法特别有吸引力，因为这本质上是在他的论文(与玻恩和海森伯合作，1926)中由他引入到量子物理学中的东西：把量子假说应用于(麦克斯韦)场本身。约当—维格纳推广的主要任务之一是，使得以费米子为量子的经典场量子化。这里我们到达了与我们的讨论相关的关键点。坦率地讲，自然界中根本没有这种经典场。当然，也没有诸如费米子这样的经典粒子。但这里的问题是，在量子场论的框架中，如何从某种假定的本体论中推导出费米子。面对这种情形，约当和维格纳能挑选出作为他们的量子场论起点的唯一候选者，就是单费米子的三维薛定谔场。

从这样一种“经典场”出发，多费米子问题的量子场论表述，能通过一种形式上与狄拉克在玻色子情形中的相似程序获得。仅有的不同在于对量子化条件的定义，它必需满足泡利不相容原理(Pauli exclusion principle)。[8] 在这种表述中，为了提高或减少某一量子态中的粒子数，引入了产生和湮灭算符(即 a，a^* 或 b，b^*)。场振幅被表达为这些算符的线性组合。这是被分解为振子振幅的电磁场量子化的直接推广。一个振子振幅的算符或创生或消灭该振子的一个量 [168]

子。因此，费米子被看作一个场量子，就像光量子是电磁场的量子一样。这种量子能被创生或消灭，因此不是永恒的或长久的。

在一次访谈中，维格纳在回顾他与约当的合作时说："正如我们通过电磁场的量子化得到光子一样，通过薛定谔场的量子化我们应该能够得到物质粒子。"(Wigner，1963)这里可以引出关于这个深层思想的一个重要问题。可设想的是，光子从电磁场中激发是由于电磁场具有物质性(即携带能量和动量)。但薛定谔场的情形如何呢？它也是一个物质场吗？或仅仅是一个概率场？罗森菲尔德(Leon Rosenfeld)告诉我们："在某种意义上或其他意义上，约当本人实际上比大多数人更为认真地把波函数看作概率幅。"(Rosenfeld，1963)但是，甚至约当本人也没有提供一个一致的说明，来证明他把概率幅用作一个经典场是正确的，这个经典场通常被看作是物质的，具有能量和动量。在我看来，经典场为单费米子的三维薛定谔场所替代，在世界观上引发了一次根本性的变革。

首先，实在论解释已取代了关于单费米子的薛定谔波函数的概率解释。也就是说，薛定谔波函数开始被看作物质波而不仅仅是概率波(或导波)。这种取代的理由很简单：否则，波将只与单粒子有密切关系，而与多粒子问题没有关系。一个甚至更重要的理由是：一个非物质概率场不能把物质性赋予它的量子，因此作为场量子的粒子，无法从场获得它们的物质性。于是，粒子的物质性将没有源头且变得神秘。在20世纪场论历史中一个确切的事实是，在量子场论中，从1928年约当—维格纳的论文发表以来，单费米子的薛定谔波总被看作是一个物质波场。在这一取代之后，最初描述单费米子运动的薛定谔方程变成了场方程。只有那时，与电磁场的量子化相类似的过程才被合法地采纳，多费米子问题通过运用产生和湮灭算符得 [169]

到解决，即费米子场论才得以发展。

其次，场本体已取代了粒子本体。物质粒子(费米子)不再被看作一种恒久的存在或独立的存在，而是(费米子)场的激发，是一种场量子。

根据量子力学的正统解释，量子场论中薛定谔波的实在论解释，不能被证明为是对量子力学的逻辑一致的发展，至多是被物质(费米子)和光(玻色子)之间的形式类比所要求的权宜之计。然而，它的验证实际上在于它随后产生的成功。因此，概率解释被实在论解释所替代，或者退一步说，经典场被薛定谔波所替代，标志着一种新的物质场的引入和理论发展的一个新阶段的开始。

一旦物理学家习惯于把单费米子的薛定谔波看作一种新的物质波，看作费米子场论的起点，那么作如下假定就很自然但却是误导的：玻色子场论的起点也是一种薛定谔波(即单玻色子的薛定谔波)，而不是经典(麦克斯韦)场，以至于两种情形都能置入一个单一的“二次量子化”的概念框架之中。

由于场的量子化是把场变量转化为满足某些代数关系 $[b_r, b_s^*] = \delta_{rs}$ 等等的算符[9]，因此表达为这些算符的线性组合的场振幅也变成算符，并满足相同的代数关系。在狄拉克(1927b, c)、约当和奥斯卡·克莱因(1927)以及约当和维格纳(1928)的理论中，算符 b_r 和 b_r^* 仅仅起了引起粒子从一个量子态跃迁到另一个量子态的作用。然而，一旦这些算符被定义，它们允许自然地推广到粒子实际上能被产生和湮灭的情形，即那时粒子总数不再是一个运动常数(见 7.4 节)。这实际上就是实验观测到的和在相互作用的相对论性量子场论中加以理论化的情形。例如，在电磁场中，光量子能被发射和吸收，因此其他玻色子，例如 π 介子，也能被发射和吸收。在电子情形中，我们知道电子—正电子对能被产生和湮灭。所有这些过程能够用场论

的产生和湮灭算符方便而精确地加以描述。这些算符也用来简洁而[170]精确地描述费米子和玻色子之间的基本相互作用，它们由两个费米子产生算符和/或湮灭算符与一个玻色子算符的乘积所组成。这被解释为一个费米子态的改变，即在一个态湮灭，在另一个态产生，同时伴随着一个玻色子的发射或吸收。关于相互作用的更多阐述将在7.5节和第三篇给出。

作为一个场，量子场和经典场一样，一定是一个连续的充盈空间。然而，量子场论中处理的是所谓的定域场(local field)。也就是说，场变量只在一个空间点被定义(描述物理条件)。正如我们在上面已强调的，定域场不是一个c数，而是一个算符——定域场算符。当量子场论刚发明时，按照物理粒子的发射和吸收(产生和湮灭)的观点，算符场有清楚的和直接的物理解释。然而，在狄拉克引入他的真空观念(1930a)之后，算符场变成了抽象的动力学变量，人们借助于这些动力学变量建构了物理态。它们本身没有任何基本的物理解释。关于狄拉克真空的更多说明将在7.4节给出。

狄拉克借助其量子场论(在该理论中，粒子被看作场的激发)，声称已证明，在波动观点和光量子观点之间存在"完全的和谐"或"完全的调和"。这一主张从提出那时起，就被科学共同体，包括大多数物理学家、物理学史学家和物理哲学家认为是正确的。例如，约斯特断言，"辐射的波粒二象性之谜，从1900年以来就强有力地推动了理论物理学的发展"，已被狄拉克1927年的论文"所解决"(Jost, 1972)。然而，这一主张已受到雷德黑德的挑战(1983)，雷德黑德的论据可以总结如下：如果场振幅是对角的，那么这种表示法可以叫做场表象(field representation)；同样，如果数字算符全部是对角的，那么这种表示法就可以叫做粒子表象(particle representation)。然

而，在后一种表象中，场振幅不具有明晰性，因为数字算符与产生和湮灭算符不对易。因此，人们不能同时把明晰的值既归于粒子数又归于场振幅，从而再次显示出波粒二象性，虽然是以不同的形式。

在上面的讨论中，狄拉克的主张也受到区分能量子和光量子以及场运动的量子化和物质场的量子化的挑战。对付后一种挑战，可以通过指出由场算符激发的场量子不是能量子或量子化的能量，而是由场的能量、动量、电荷和所有其他属性的量子化载体来处理。在这种情形中，场量子和爱因斯坦的光量子(或其他玻色子或费米子)之间的区别消失了，并且场系统能按照粒子或场量子来描述。这是真 [171] 的。但这不是狄拉克的观点。在狄拉克1927年的论文中，为场算符所激发的是场的量子化能量(能量子)，而不是场能的量子化载体。因此，从本体论上讲，把场的量子化能量(动量、电荷等)与场的能量(动量、电荷等)的量子化载体视为同一，仍是一个大问题。

应该顺便提到，电磁场的能量和动量不是通常的四维矢量形式。它们是四维张量的分量，其行为与粒子的相应的量很不相同。因此，按照其能量、动量的洛伦兹变换，场通常没有粒子的属性。然而，正如费米所表明的，场可以被分成一个横向部分(光波)和一个纵向部分(静态库仑场)。前者在洛伦兹变换下的行为与粒子有某种相似性，因为它的能量和动量形成一个四维矢量。但对后者来说，情形并非如此。在费米量子理论(1932)中，只有前者服从量子化，导致光量子的存在(它在其他某些方面的行为也像粒子)，而后者仍然未被量子化。

诚然，对场进行量子化可以使场的粒子方面显现出来。但是，量子场论毫无疑问要比从它的粒子解释中可看见的东西更为复杂和丰富。所谓零点涨落(zero point fluctuations)的存在，暗示着甚至在真

空中(即在没有粒子或场的量子出现的态中)场也存在。涨落的真空场实际上产生了可观测的效应。例如，我们可以把自发辐射看作在真空涨落影响下发生的受迫发射。另一个效应是空间中电荷位置的涨落，它产生了粒子的部分自能(Weisskopf，1939；另见 7.6 节)。束缚电子的束缚能由于这些位置涨落而减少，这解释了兰姆效应中的大部分水平移位(Welton，1948；Weisskopf，1949)。在费米子场情形中，在某些有限时空体积中，当粒子数一定时，也有电荷和电流密度的涨落。

狄拉克没有解决场的波粒二象性之谜，因为他没有找到连结连续场和离散粒子之间的桥梁。事实上，他只是通过场的量子化表述和他对这一表述的特殊解释，把场还原为粒子。他的解释最特别之处在于其零态光量子的概念，这一概念表明，对他来说，光量子是一种 [172] 稳定存在的真实粒子(他把光量子的产生和湮灭仅仅看作跃出和跃进零态的真实过程的显现)。简言之，对狄拉克而言，场的量子化意味着场由离散和稳定的粒子或量子组成。这是粒子本体论解释赖以发展的基础。

事实上，解决这个谜的关键一步是由约当完成的。约当的数学表述与狄拉克的相同。两者的不同在于表述的解释所支撑的本体论承诺。狄拉克从真实的连续电磁场出发，但却把它还原为一组离散的粒子。相反，约当工作的起点是薛定谔波。他不但在 1928 年处理费米子场的论文中，而且在 1927 年的论文中，也把电磁场看作单光量子的薛定谔波函数。

我们在上文中已注意到，约当很认真地考虑了薛定谔波的物理实在性。这个立场与他对量子物理实在的本质和波粒二象性的互补性理解密切相关。根据约当：

量子物理实在以一种显著的方式，比经典理论所寻求的表征它的观念体系显得简单。在经典表征体系中，类波辐射和微粒辐射是两种根本不同的东西；然而，实际上，只有单一类型的一种辐射，两种经典表征都只给出它的一部分正确的图景。

(1928)

也就是说，真实的情况是，辐射的类波和微粒两个方面，应看作同一种基本物质的不同的特殊显示。这里，连接连续场和离散粒子的桥梁是：作为量子物理实在的辐射既是波也是粒子。

在经典理论中，粒子由于它们稳定的存在，根本不同于场的能量子，因为场的能量子能被产生和吸收。在量子理论中，根据约当，考虑到实在既是波又是粒子，粒子也能被产生或吸收，正如波场的量子一样；而场也能展示它们的离散存在。以这种方式，波粒二象性之谜似乎已为约当而不是为狄拉克所解决。

有趣的是，这一谜得以解决所植根的基本本体是某种既不同于经典粒子，也不同于场的新东西。它不能被看作经典粒子，因为从它跃出的量子没有稳定的存在和个体性(individuality)。它也不能被看作经典场，因为它作为量子化的场，已失去了连续性。一个被雷德黑德称作朝生暮死的新的本体已出现，据此，量子场论随后得以发展。 [173]

7.4 真空

算符场和场量子化，特别是约当—维格纳量子化，只有用真空态的说法才有其直接的物理解释。为了对此有清楚的了解，让我们来考察一下20世纪20年代晚期物理学家引入场量子化前后的真空

概念。

爱因斯坦终结了以太概念之后，无场和无物质的真空被认为是真正空的空间。然而，这种情形随着量子力学的引入而发生改变。从那时起，真空再次变得充盈。在量子力学中，对于光量子数目 N 和场振幅的相位 θ 来说，不确定关系 $\Delta N \Delta \theta \geqslant 1$，意味着如果 N 有一个给定的零值，那么场将显示它对于平均值零的某种涨落。[10]

充盈真空的下一步由狄拉克作出。在他关于电子的相对论性理论(1928a，b)中，狄拉克遇到了一个严重的困难，即负动力学能态的存在。作为对这个困难的一个可能的解决方案，他提出一种新的真空概念："所有的负能态都已被占满"，而所有的正能态没有被占满(Dirac，1930a)。因而，由于泡利不相容原理，一个电子从正能态到负能态的跃迁不可能发生。这个真空态不是一个空的态，而是一个充满负能电子的海洋。这个负能电子海作为一个普适的背景是不可观测的，然而在负能海中的空穴则是可观测的，这些空穴的行为就好像一个具有正能和正电荷的粒子。狄拉克最初试图把空穴看作质子。这一解释不久被证明是不成立的。1931 年 5 月，狄拉克接受了奥本海默(1930b)和外尔(1931)的批评，把空穴看作一种带有正电荷和电子质量的新粒子(Dirac，1931)。

显而易见，狄拉克在 1930 年到 1931 年间的不可观测的无穷多负能电子海，类似于他 1927 年的不可观测的零能量光子的无穷集。这样一个由不可观测粒子组成的真空，允许粒子的产生和湮灭现象的存在。[174] 假定有足够的能量，一个负能电子会上升到一个正能态，对应于一个正电子(负能海中的空穴)和一个普通电子的产生。当然，相反的湮灭过程也可能发生。

人们可能会说，如温伯格(1977)所做的，如果没有引入量子场论

思想的话，那么所有这些过程能用可观测态和不可观测态之间的跃迁概念得以说明。这是真的。但是另一方面，准确地说，真空和跃迁的概念已为量子场论提供了一个本体论基础。事实上，负能海的引入，作为狄拉克关于电子的相对论性理论所造成的困难的一个解决方案，隐含着如果不涉及一个无穷多粒子系统，那么单电子的一致的相对论性理论将是不可能的；也隐含着量子场论描述对于相对论性问题（下面将有更多的讨论）是必要的。此外，可观测态和不可观测态之间的跃迁概念，的确为量子场论中关于激发的思想提供了一个原型，因此给出了场算符的一种直接的物理解释。

我们注意到，产生和湮灭的概念在时间上先于量子力学，可以追溯到20世纪之初，其时这些概念建立在物质和光的以太模型基础之上(Bromberg，1976)。事实上，狄拉克的真空观念作为一种基质，具有与以太模型相同的一些典型特征，这就解释了为什么狄拉克在晚年又回到了以太观念[11]，尽管在最初对产生和湮灭过程的描述中，他并没有明确诉诸以太模型。

狄拉克真空与以太相类似的一个显著特征是，它的行为就好像一个可极化的介质。一个外在的电磁场使负能海的单电子波函数发生畸变，因此产生了一个电荷—电流分布，起着对抗施感场的作用。结果，粒子的电荷表现为减少。也就是说，真空能被电磁场极化。电荷和电流（它们甚至出现于无电子的真空态）的涨落密度，能被看作电磁场中涨落的电子—正电子场的相对物。

总之，引入充盈真空后的情形就是如此。假若我们从一个电子—正电子场 Ψ 开始，它将产生一个对初始场 Ψ 起反应并改变它的伴随电磁场。同样，一个电磁场将激发此电子—正电子场 Ψ，相关的电流作用并改变着初始电磁场。因此，电磁场和电子—正电子场密

[175] 切相关，它们中没有一个具有独立于另一个的物理意义。因而，我们拥有的是一个由电磁场和电子—正电子场组成的耦合系统，并且对一个物理粒子的描述将不会被先验地写下，而只会在解决了一个复杂的动力学问题之后才出现。

真空实际上是存在着的无穷多负能电子猛烈而杂乱无章地活动（涨落）的舞台，这个思想后来通过排除这些电子的真实出现的观念而得以缓和。正是奥本海默和弗里（Oppenheimer and Wendell Furry，1934）意识到，通过协调一致地交换那些产生和湮灭算符所扮演的角色（它们对负能态起作用），可以放弃充盈真空的假设，而不对狄拉克方程有任何根本性的改变。用这种方式，作为单粒子的两个可供选择的态，电子和正电子对称地进入形式体系，真空的无穷大电荷密度和负能密度消失了。用这个形式体系，真空再次变成一个明显没有粒子的物理上合理的态。最重要的是，在相对论性理论中，真空的能量和动量必须为零，否则这个理论的洛伦兹不变性将不能保持。

同年，泡利和魏斯科普夫（Pauli and Weisskopf，1934）在他们所从事的对标量粒子的克莱因—戈尔登相对论性波动方程进行量子化的工作中，使用了为负能态交换产生和湮灭算符的相同方法。标量场的量子理论包含空穴理论的所有优点（粒子和反粒子，粒子对的产生和湮灭过程，等等），而没有引入一个充满粒子的真空。

总之，上述这两篇论文表明，量子场论能够自然地归并反物质的思想，而不用引入任何不可观测的负能粒子。因此，量子场论能够令人满意地描述粒子和反粒子的产生和湮灭，粒子和反粒子现在被看作地位平等的场量子。

对于大多数物理学家来说，这些进展，特别是温策尔有影响力的著作《场的量子理论》（*Quantum Theory of Field*，1943）的出版，已解

决了这件事：一个无穷大负能电子海的图景已被看作物理学史上的古玩，并且已被遗忘(Schwinger，1973b；Weinberg，1977；Weisskopf，1983)。然而，仍有一些人坚持认为狄拉克的真空观念具有革命性的意义，其真空不空而是充满物质的思想精华仍然保留在我们现在的真空概念之中。[12]他们争辩说，充满物质的真空概念得到了如下事实的强有力支持：真空中物质密度的涨落，作为真空除了电磁真空涨落之外的一个额外属性，甚至在负能电子消除后仍然保持。

这里，我们遇到了量子场论中最深刻的本体论困境。一方面，[176]根据狭义相对论，真空必定是零能量、零动量、零角动量、零电荷，无论什么都为零的洛伦兹不变态，也就是说，是一个虚空态。如果考虑到在现代物理学和现代形而上学中，能量和动量已被认为是物质的本质属性，那么很明确，真空不能看作一种物质。另一方面，真空中存在的涨落强烈地表明真空必定是有物质的东西，当然不是虚空。走出这一困境的可能途径之一可以是：重新定义“物质”，去除能量和动量是物质的确定属性的规定。但是，这也将是一个特设，不可能从其他例子中找到支持。另一个可能途径是把真空看作一种前实体(pre-substance)，一种具有潜在物质性的底层基质。如果有一些其他属性也被注入其中，真空能被能量和动量激发而变成物质，它将变成物理实在。

无论如何，没有人能否定真空是不存在真实粒子的态。只有当我们扰动了真空，只有当我们用能量和其他属性激发了真空，真实的粒子才会存在。对于定域场算符(其直接的物理意义被认为是一个物质粒子的激发)来说，我们能说些什么？首先，局域激发被作用于真空的一个定域场算符 $O(x)$ 所描述，意味着能量、动量和其他特殊属性在一个时空点上注入了真空。这也意味着，由于不确定性关系，

对于各种物理过程，任意数量的能量和动量都可利用。显然，由 $O(x)|$真空$\rangle$ 符号表现的这些属性的物理实现，将不仅仅是一个单粒子态，而且必定是所有合适的多粒子态的一个叠加。例如，$\Psi_{\mathrm{el}}(x)$ $|$ 真空$\rangle$ $= a\,|$一个电子$\rangle$ $+ \sum a'\,|$一个电子 + 一个光子$\rangle$ $+ \sum a''\,|$一个电子 + 一个正电子 + 一个电子$\rangle$ $+ \cdots$，其中 $\Psi_{\mathrm{el}}(x)$ 是所谓着衣场（dressed field）算符，a^2 是 $\Psi_{\mathrm{el}}(x)$ 激发形成的单个裸粒子态的相对概率，等等。

结果，场算符不再指物质粒子，而是变成抽象的动力学变量，借助于它，人们构建了物理态。那么，我们如何从底层动力学变量（定域场算符）这个理论的起点，到达可观测的粒子呢？这是一个任务，它暂时只有在重正化（renormalization）程序帮助下才得以实现。7.6 节和 8.8 节将更详细地讨论重正化。

[177] 7.5 相互作用

粒子穿越空间作用于另一粒子的相互作用机制，是物理学中的一个基本问题，以至于整个场论纲领能被看作是对它所做的一个回答。在牛顿万有引力定律表述中，物体不论相距多远，都可以不借助于介质，瞬时地相互作用；与此表述形成对照的是，在电磁现象领域，法拉第和麦克斯韦发展了力场的思想，在他们的理论中包含了介质的引入，相互作用体之间的力是通过介质而被逐次、连续的传播所传输。但是这样一种力场的思想遇到了两个致命的困难，一个是物理上的，一个是概念上的。从物理上讲，人们不能在一个力学模型的形式中为介质思想找到一个令人信服的基础，而又不会产生与光学证据相矛盾的结果。从概念上讲，这一观念导致一个无限的递归：每个力场都需要一种介质，以使它能在与它匹配的粒子间传播，这种介质的构

成粒子又必须为另外一个场所匹配，这要求另一介质来实现它的传播，如此等等。

为了规避这一概念上的困难，赫兹把粒子间的耦合描述为由一个未知的“中间项”实现的传递过程：

> 第一个物体的运动确定了一个力，然后这个力确定了第二个物体的运动。这样，力能以平等的合法性被看作总是运动的原因，同时也是运动的结果。严格地讲，这是只在两个运动间设想的一个中间项。
>
> (1894)

赫兹的思想在场论历史中是重要的，因为，第一，它把相互作用从一个普适介质中解脱出来；第二，它允许相互作用传递子的存在。但是它仍然遗留下重大的问题：神秘的“中间项”是什么？回答这个问题的决定性进展由爱因斯坦作出，爱因斯坦认为赫兹所发现的电磁场就是基本的实体，并把它看作在粒子间传递相互作用的介质。值得注意的是，爱因斯坦关于电磁场的思想比赫兹的“中间项”思想更具有普遍性，因为它能独立存在，而赫兹的“中间项”的存在依赖于相互作用粒子的存在。

狄拉克在发展量子场论的过程中，遵循了赫兹的思想和爱因斯坦的思想，把相对论性经典电动力学中相互作用的观点总结为“每个粒子都发射波，这些波以有限速度向外传播并且影响穿过它们的其他粒子的思想”，他声称“我们必须找到一种把新的(相对论性的)信息接收进量子理论的方式”(Dirac，1932)。 [178]

事实上，带电粒子通过电磁场相互联系、电磁场由相互作用的粒

子发射和吸收这一经典观念，的确已作为一种概念模型而被接收进量子场论的相互作用概念之中。经典观点和量子观点之间的不同在于两点。其一，根据量子观点，相互作用由电磁场的离散量子传递，这些相互作用"连续地"从一个粒子经过另一个粒子[13]，而不是通过一个同样连续的电磁场传递。因此，在量子理论中，离散粒子和连续力场之间的经典区分是模糊的。其二，从量子观点看，相互作用粒子一定被看作费米子场的量子，正如光子是电磁场的量子。所有这一类量子，不同于经典的不变粒子，它们能够产生和湮灭。因此，不但在传递粒子间相互作用的介质是一个场的意义上，而且也在相互作用粒子本身应该被看作场即费米子场的一种表现形式的意义上，这是一个场论。这样一种费米子场能通过它的量子传递相互作用，就好像电磁场通过光子传递相互作用一样。因此，物质和力场之间的差别从这一场景中消失了，并且被一种普适的粒子—场二象性取代，这种粒子—场二象性同等地影响每一个构成实体。

量子电动力学建立在约当(和玻恩合作，1925；和玻恩、海森伯合作，1926)、狄拉克(1927b；c；1928a，b；1930a；1931)、约当和维格纳(1928)、海森伯和泡利(1929；1930)工作的基础之上[14]，它是关于相互作用的量子理论沿着两条路线发展的起点。一条路线导致了费米的 β 衰变理论，即弱相互作用的量子理论的原型，也导致了汤川秀澍(Hideki Yukawa)的介子理论，即强相互作用的量子理论的原型。另一条路线导致了规范理论。这一节的剩余部分将讨论第一个发展，而第二条路线将在第三篇讨论。

相互作用的量子理论的早期发展与核物理学密切地纠缠在一起。事实上，核 β 衰变和稳定原子核的核力是两个主题，它们对于相互作用的新思想的发展提供了主要的促进因素。

1932年，查德威克(James Chadwick)发现了中子，被公认为核物理学史上的分水岭。在此之前，人们接受的核结构的理论是原子核 [179] 的电子—质子模型；根据这一模型，原子核中除了质子外，还包含电子，因此在原子核中额外的电荷能被抵消，β衰变能得到解释。这个模型遇到了严重的自旋统计困难(spin-statistics difficulty)，这一困难来自拉塞蒂(Franco Rasetti)对 N_2 的拉曼光谱的测量(1929)，它暗示着由21个自旋为1/2的粒子(14个质子和7个电子)组成的氮原子核服从玻色统计。因为，正如海特勒和赫茨贝格(Walter Heitler and Gerhard Herzberg, 1929)指出的，拉塞蒂的分析意味着，如果电子—质子模型和维格纳规则[15]都正确的话，那么核内电子及其自旋将丧失对原子核的统计决定作用。

作为对这一困难的回应，苏联物理学家安巴楚勉和伊凡宁柯(V. Ambarzumian and D. Iwanenko, 1930)提出，原子核中的电子“失去了它们的个体性”，就好像原子中的质子一样，而β发射与原子的光子发射相似。这一观点的新颖之处在于电子被产生和湮灭的可能性，因此总电子数未必是一个常数。[16]应该注意，这个观点之所以是新颖的，是因为它不同于狄拉克基于空穴理论的电子—正电子对的产生和湮灭的观点，狄拉克的观点只是一个电子在一个负能态和一个正能态之间的量子跃迁，总电子数是守恒的。它也不同于约当和维格纳的工作(1928)；在约当和维格纳的理论中，产生和湮灭算符在处理非相对论性多粒子问题中，只是描述不同量子态之间量子跃迁的数学工具，并没有改变总电子数。这个新颖观点提前3年预见了费米β衰变理论的一个重要特征。事实上，这一观点被费米(1933, 1934)明确地用作建立他的β衰变理论的三个主要假设之一。

查德威克对中子的发现(1932)给电子—质子模型以致命的打击，并为原子核结构的中子—质子模型的建立开辟了道路。伊凡宁柯(1932a，b)又是把中子看作一个自旋为 1/2 的基本粒子并提出中子—质子模型的第一个物理学家。几乎在同年，海森伯(1932a；b；1933)提出了与伊凡宁柯相同的模型，并且强调(1932a)中子作为一个基本粒子将解决“自旋和统计”困难。他也建议中子和质子可以被看作同一个重粒子的两个内部态，由一个位形波函数描述，在这一位形波函数中必定包含重粒子的一个内禀坐标 ρ 。这个内禀坐标 ρ 可假设只有两个值：+1 为中子态，−1 为质子态。ρ 坐标的引入显示了海森伯模型的非相对论本质，它与量子场论没有任何关系。然而，它预见了 6 年之后作为一个内部(非时空的)自由度的同位旋(isospin)概念。同位旋的概念由凯默(Nicholas Kemmer)在 1938 年引入，并在随后的核力理论和规范理论的发展中起着重要的作用。[17]

[180]

然而，荒谬的是，海森伯试图在上述模型中合并中子概念，中子被描绘成一个由质子和电子构成的紧紧束缚的复合体，在这个复合体中，电子丧失了大部分特性，特别是它的自旋和费米子特征。海森伯相信，中子作为一个复合粒子，在核内衰变为一个质子和一个电子。这将能解释 β 衰变。在海森伯有关中子的复合模型的思想中更重要的是，他相信中子通过交换它的组分电子而与质子紧密结合，类比于在 H_2^+ 离子的化学键中电子的交换。因此，核力被认为是由“赝”电子的交换形成的。

海森伯交换力的基本思想为马约拉纳(Ettore Majorana，1933)所接受，不过马约拉纳抛弃了与化学键的类比。马约拉纳从已观测到的原子核的性质出发，特别是从核力的饱和性(saturation)和 α 粒子的特殊稳定性出发，认识到他能通过一种不同于海森伯的交换力而实现

核力的饱和。他舍弃了海森伯绘景(在海森伯绘景中，中子和质子处于两个不同的位置，它们只交换电荷而不交换自旋)，消除了海森伯的 ρ 坐标，并且不仅涉及电荷的交换而且也涉及自旋的交换。这样，马约拉纳发现核力在 α 粒子上是饱和的，而海森伯的模型错误地使核力在氘核上达到饱和。

另一个重要的发展从 β 衰变的研究中获得了更直接的激励。查德威克发现了 β 射线具有连续能谱(1914)，这只允许有两种可能的解释。一种解释为玻尔(1930)所倡导，玻尔说能量守恒在产生 β 辐射的相互作用中只是统计地有效。另一种解释由泡利提出：

> 在原子核中可能存在电中性粒子，我称之为中子[18]，它具有1/2自旋，遵守不相容原理……通过假设在 β 衰变中一个中子随着电子一起发射，连续的 β 能谱就变得可以理解，以这种方式，中子和电子的能量之和是常量(1930)，[或]等于与 β 能谱的上限相对应的能量。
>
> (1933)

在泡利最初的提议中，中微子* 被认为是原子核的构成要素，在发射之前就已存在于原子核中。中微子能在它发射的瞬间和电子一同被创生的思想，首先由佩林(Francis Perring，1933)清楚地提出。佩林的思想与伊凡宁柯和安巴楚勉的思想一起，被费米吸收从而形成其 β 衰变理论，费米 β 衰变理论作为对先前发展的一个综合，至今在许多方面仍然是相互作用的标准理论。 [181]

* 当时泡利称为“中子”。——译者

首先，费米的理论清楚地表达了相互作用的过程不是跃迁的过程、而是粒子产生和湮灭过程的思想：

在β发射发生之前，电子的确不像所想的那样存在于原子核中，但当它们被发射时，它们瞬间就获得了它们的存在；以相同于光量子的方式，它们在量子跃迁中为原子所发射……因此，电子的总数和中微子的总数（和辐射理论中光量子的总数一样）将不是恒定不变的常数，因为可能存在那些轻粒子产生或湮灭的过程。

(1933)

其次，这一理论假定狄拉克和约当量子化概率幅的方法，能被用来处理场量子（电子、中微子）数目的变化：

电子的概率幅 ψ 和中微子的概率幅 ϕ，以及它们的复共轭 ψ^* 和 ϕ^*，被认为是作用于电子和中微子量子态占有数的函数的非对易算符。

（同上）

费米对狄拉克—约当量子化的解释以及对这一量子化的场论应用，很快就成为量子场论中的正统解释；在场论符号中，费米的表示法成为标准的表示法，这一表示法保留了有关概率幅的量子化的物理解释的开放性。这就是为什么费米β衰变理论被魏斯科普夫(1972)看作“现代场论第一个范例”的原因。

再次，费米的四线相互作用(four-line interactions)表示法，在20世纪70年代初之前一直支配着弱相互作用理论，它作为一个有效近

似，现在仍被用在低能弱相互作用领域：

> 一个中子转换为一个质子，必然伴随着一个电子和一个中微子的产生，这个电子被观测为一个β粒子；反之，一个质子转换为一个中子，必然伴随着一个电子和一个中微子的湮灭。
>
> （同上）

总体上讲，费米的弱相互作用理论是在类比量子电动力学的基础 [182] 上建立起来的。不但电子和中微子的产生和湮灭是通过类比光子的产生和湮灭而假定的，而且哈密顿函数中的相互作用项也是采用类似于量子电动力学中的库仑项的形式(Fermi，1934)。这样一种类比也为汤川秀澍、杨振宁和米尔斯(Robert Mills)所采用，并且也为许多希望把现存理论扩展到一个新领域的其他物理学家所采用。

无论费米β衰变理论可能是多么地重要，然而，他完全没有触及核力与核场的量子。就重子(质子和中子)本身而言，与通过“二次量子化”方法处理的轻子(电子和中微子)形成对照，它们被费米(就像海森伯一样)视为重子的两个内部量子态，由“位形空间中的通常表示法”描述，而“重子的内禀坐标ρ必定包含在位形空间之中”(同上)。

显然，重子在这里被看作是非相对论性粒子而不是场量子。因此，对于重子来说，没有产生和湮灭过程，只有两个态之间的跃迁。从这方面来看，费米的理论似乎只是一个过渡性的理论，而不是一个成熟和自洽的表述。由于这个原因，人们倾向于通过把质子和中子也看作场量子而把它改造为一个自洽的量子场论表述，或者以一致的方式回到跃迁的老观念，把电子和中子的产生和湮灭看作量子跃迁的结果或显现。后一条路线被所谓的费米场模型(Fermi-field model)采

纳。根据这一模型，电子和中微子的同时产生被看作一个轻子从一个中微子的负能态跃迁到一个电子的正能态。[19]这里，我们能清楚地看到狄拉克充盈真空的思想对于量子场论随后的发展的深刻影响。

费米β衰变理论也隐含着从一个电子—中微子对的交换演绎出中子和质子之间的交换力的可能性。这种可能性由塔姆(Igor Tamm, 1934)和伊凡宁柯(1934)发展的所谓"核力的β理论"或"核力的[电子—中微子]对理论"作了探究。在这个理论中，一个质子和一个中子能通过虚发射和再吸收一个电子—中微子对而相互作用。也就是说，核力与β过程相关。但是，作为一种二阶效应，这个相互作用非常弱。因此，甚至是这个模型的提出者，也都对核力的起源是否在于β衰变和它的逆过程，表示了恰当的怀疑。

[183] 核力的[电子—中微子]对理论是短命的，但在历史上却是重要的。作出这一论断的理由是双重的。其一，它在物理学史上第一次证明：相互作用能通过有限(即非零)质量的量子而传递。其二，这一理论澄清了，核相互作用的短程性是由传递核力的量子拥有非零质量这一事实所直接决定。它还给出了估计核力场量子质量的一般方法的线索。例如，塔姆在1934年指出，交换能依赖于相互作用路程 r 的减函数 $I(r)$，当 $r \ll h/mc$ 时，$I(r) = 1$，其中 m 是传递相互作用的电子的质量(1934)。

从理论上讲，已知"中子和质子的相互作用是在 $r \sim 10^{-13}$ cm 的量级上"，塔姆就应该能够推断出传递相互作用的粒子的质量大约为 200 m_e(m_e 是电子质量)，就像汤川秀澍后来所做的那样。从技术上讲，由于这里参量 m 已被预设为是电子质量，因此这样一种马后炮是无意义的，用另一种传递核力的粒子取代电子是将要采取的一大

举措。然而，核力的粒子对交换理论在某种意义上已为采取这样一个关键性举措铺平了道路。

在核力理论或广义相互作用理论的发展过程中，汤川秀澍“关于基本粒子的相互作用”(1935)或简称汤川介子理论的工作，是转折点。汤川于1933年开始他的研究，其时他受海森伯原子核和复合中子模型的驱动，曾企图通过在质子和中子之间交换一个电子来解释核力。在这一不成功的尝试和所有其他相似的尝试中，一个众所周知的内在困难是自旋和统计的不守恒。幸运的是，仁科芳雄(Yoshio Nishina)给了汤川一个好的建议，即交换一个玻色子能克服这一困难。[20]

之后不久，费米、塔姆和伊凡宁柯的论文发表。交换力的观念，粒子作为场量子的产生和湮灭的观念，非零质量的场量子传递核力的观念，对汤川来说，都是可利用的。核力不可能直接与β过程相联系，这一认识对汤川来说也是可利用的。于是，在1934年10月初，汤川构想出一个创造性的思想：一个与重子有强相互作用且具有约200 m_e 质量的玻色子，可能是核力场的量子。这一思想是其介子理论的基础，也是整个现代基本相互作用理论的基础。

[184] 汤川思想的新颖之处在于以下几点。一、核力由一个力场，即一个玻色子场传递，或者说由单个量子而不是由一对粒子传递。二、核力是一种相互作用，这种相互作用不同于β过程的相互作用，因为“中子和质子之间的相互作用比起费米情形中的要大得多”(Yukawa，1935)。这一思想是强相互作用和弱相互作用之不同的由来，也是量子场论的基本原则之一。三、核场或者说“U场”，作为一个新的物理客体，不可还原为已知的场。这一点把汤川关于核力所做的工作与所有以前的工作区分开来，也是冲破20世纪20年代西

方物理学思想禁锢的一步，按照当时的物理学思想，人们不应该增加“不必要的”实体。自汤川的工作以后，物理学的主导思想是：自然界中的每一种基本相互作用，都应该由一个中间玻色子场刻画。

有趣的是，我们注意到，汤川也提出玻色子可能是β不稳定的，因此原子核的β衰变可由一个两步过程来解释，这个过程由一个虚玻色子传递。用这种方式，汤川的玻色子扮演着双重角色，既是核力量子，又是处于虚态的β衰变中间物。虽然汤川的方案不得不分成两部分[21]，但是理论的这样一种统一的特征的确很有魅力。自20世纪50年代后期以来，汤川关于弱相互作用的三线表示法(一条线代表中间玻色子，另两条线代表费米子，它们相互耦合)再次变得流行。确实，在后来的理论中，作为传递弱相互作用介质的玻色子，不同于作为传递强相互作用介质的玻色子。然而，在表述基本相互作用的两种情形中，基本思想是相同的，正如汤川所表明的，它本质上是以电磁相互作用为模型构造的。对于物理学共同体的大多数人来说，这样一种形式上的相似性似乎表明，在基本相互作用中存在着一个更深层次的统一。

在非阿贝尔规范理论(non-Abelian gauge theories)登上历史舞台之前，在量子场论中，潜藏在相互作用概念之下的思想可以总结如下。场量子(费米子或玻色子)之间的相互作用由另一种场量子传递，各种量子都可以产生和湮灭。由于作为定域场的介质场有无穷大的自由度，在计算相互作用时不可避免会发散。这是量子场论固有的一个致命缺陷，它深深地植根于相互作用的基本思想之中。因此，如果不解决这个严重的困难，量子场论将不能被看作是一致的理论。这里，我们遭遇到了重正化问题，重正化在量子场论的概念结构中起着重要作用，7.6节、8.8节和11.4节将讨论这一问题。

7.6 重正化

[185]

无穷大的经典根源

在早期量子场论中，发散困难(divergence difficulty)与电子的电磁自能中出现的无穷大(infinities)紧密相关。然而，这不是量子场论特有的新问题，因为在经典电动力学中，这一问题就已出现了。根据J·J·汤姆孙(1881)，包含在一个半径为 a 的球形电荷场中的能量与 $e^2/2a$ 成正比。因此，当一个洛伦兹电子的半径趋于零时，能量线性地发散。但是，如果电子被赋予一个有限的半径，那么在电子球中库仑斥力将使组态不稳定。庞加莱(1906)对这个佯谬的反应是，认为在电子中可能存在一个用来平衡库仑力的非电磁内聚力，这样电子就不会不稳定。这一模型的两个原理对后来几代人都产生了巨大的影响。第一，存在这样一种观念，认为电子的质量至少部分有一个非电磁起源。第二，非电磁的补充相互作用与电磁相互作用结合在一起时，就导致了电子的可观测质量。例如，斯蒂克尔伯格(Ernest Stueckelberg，1938)、波普(Fritz Bopp，1940)、派斯(Abraham Pais，1945)、坂田昌一(Shoichi Sakata，1947)和许多其他人，在他们研究电子自能问题时，都从庞加莱的思想中获取灵感。

正如费米1922年首先指出的(参阅 Rohrlich，1973)，庞加莱电子的平衡是不稳定的，而且是扭曲变形的；这一看法引出对这一困难的另一种反应，由弗伦克尔(Yakov Frenkel)在1925年首先提出。弗伦克尔在经典框架内论证道：

> 一个广延电子的内部平衡……从电动力学的观点看，变成了……一个不可解决的谜。我认为这个谜(及与它相关的问题)是一个故弄玄虚的问题。它的产生源自把划分原则不加批判地应用

于物质的基本组成成分(电子)，当把这一划分原则应用于复合系统(原子等)时，才会得出这些确实是“最小的粒子”。这些电子不但物理上不可分，而且在几何上也不可分。它们在空间中完全没有广延性。电子的组分之间的内力不存在，是因为这样一些组分无法获取。质量的电磁解释因此被排除；并且与在洛伦兹理论的基础上确定一个电子精确的运动方程有关的所有那些困难也一起消失了。

(Frenkel, 1925)

[186] 弗伦克尔关于点电子的思想迅速为物理学家所接受，并成为量子场论中局域激发和局部相互作用思想的一个概念基础。寻找电子的结构这一思想被放弃了，因为，如狄拉克在 1938 年的论文中提出的，“电子是一个过于简单的东西，以至于不能产生支配其结构的定律的问题”。因此，很显然，在定域性假设(locality assumption)背后隐藏着一个预设，那就是：承认我们对电子结构的无知，和对量子场论描述的其他基本实体的结构的无知。这种对于点模型和随之而来的定域性假设的正当性在于，它们在目前实验可以获得的能量上构成了近似真实的表示，毕竟目前实验可获得的能量太低，以至于无法探究这些粒子的内部结构。

通过采用点模型，弗伦克尔、狄拉克和许多其他物理学家消去了电子组分之间的“自相互作用”，因此也就取消了电子的稳定性问题。但是，在不放弃麦克斯韦理论的前提下，他们不能消去点电子和它产生的电磁场之间的“自相互作用”。这个弗伦克尔悬而未决的问题，随着量子场论的出现，变得更为尖锐了。

量子场论中的无穷大痼疾

奥本海默(1930a)注意到，当详细考虑带电粒子和辐射场之间的

耦合时，尽管最低阶的第一项不会造成任何麻烦，但微扰方法的较高项总是包含无穷大的量。事实上，在海森伯和泡利关于量子化场的相对论性理论的一般表述的论文(1929)中，这一困难就已经出现。他们发现，在不考虑负能电子的前提下，计算一个电子的自能时，在洛伦兹理论中出现的点电子的发散困难仍保留在量子场论中：虚光子的动量积分平方地发散。海森伯—泡利的方法被奥本海默用来研究原子中的能级，这就陷入了困境。为了获得原子的有限能量，就必须去掉所有的自能项。但这样一来，这个理论将不是相对论性不变的。另一方面，如果保留自能项，那么理论将导致一个荒谬的预言：谱线将无穷大地偏离由玻尔频率条件计算出的值，并且“两个不同态的能量差一般不是有限值”。这里，所有的无穷大项来自电子与它自己场的相互作用。因此奥本海默得出的结论是，发散困难使量子场论与相对论不相调和。沃勒(Ivar Waller)对孤立点电子的自能计算也导致发散结果。因此，到1930年情况已变得足够清楚，自能发散困难仍然保留在量子场论中。[187]

经典理论量子化修正的失败尝试

狄拉克、温策尔和其他一些物理学家认为，这个困难可能源于经典点电子理论，在经典点电子理论中，早已有了自能困难。温策尔(1933a，b；1934)在经典框架内把一个“限制性程序”引入到洛伦兹力的定义中。从狄拉克的多时间理论(multi-time theory；Dirac，1932)出发，根据这一理论，每个带电粒子被赋予一个独特的时间坐标 t，温策尔提议“场点”在类时方向上应该逼近于“粒子点”。通过这种方式，在经典理论中，与麦克斯韦场相互作用的点粒子能够避免自能困难。这里，所谓的“逼近”是通过引入一个密度因子 $\rho_0(k)$ 实现的，对于一个点电子来说，$\rho_0(k)$ 应当等于1。借助于

$\rho_0(k)$，静电自能的表达式可写为 $U = e^2/\pi\int_0^\infty \rho_0(k)\mathrm{d}k$。直接假设 $\rho_0(k) = 1$，静电自能将线性发散。如果人们首先把 $\rho_0(k)$写作 $\cos(k\lambda)$，并假设只在积分后 $\lambda = 0$，那么 $\rho_0(k)$仍将为1，并且所描述的仍是一个点粒子，不过 U 显然等于零。

借助于这个形式方法，人们能避免由库仑场引起的自能发散。这种方法能容易地扩展到相对论情形，但和与横场相联系的描述自能特征的发散没有关系，对于量子场论来说，这种发散是特有的。为了处理这种特有的发散，狄拉克在他 1939 年的论文中，试图通过运用温策尔的限制性程序来修改场的对易关系。然而，狄拉克在他 1942 年的论文中发现，为了实现这一目标，有必要引入一个负能光子和负概率，即所谓的不定度规(indefinite metric)。因此，上面提到的经典方案必须彻底放弃。

量子场论中的自能困难

魏斯科普夫(1934，1939)首先最详尽地研究了自能问题；他表明在基于狄拉克真空绘景的量子场论框架中，一个具有无限小半径的电子的自能，对于在以 e^2/hc 的幂展开的表达式中的一阶近似来说，只是对数地无穷大。

[188] 魏斯科普夫分析的主要步骤如下。通过与经典理论相比较，他发现，由于三个原因，单电子的量子理论已把自能问题置于一种不同的情形。第一，必须假定电子半径为零，因为如果只有一个电子出现，那么"对于点之间所有可能的有限距离来说，在两个不同的点同时找到一个电荷密度的概率为零"。因此，静电场的能量 $W_{st} = e^2/a$ 是线性地无穷大。第二，电子有一个产生一个磁场和一个交变电场的自旋。这个自旋的电场能量和磁场能量一定相等。另一方

面，在一阶近似上，部分地依赖于电子自能的电荷，由$(\pi/8)\int(E^2-H^2)\mathrm{d}r$给出。因此，如果自能用场能来表达，那么电的部分和磁的部分有相反的符号。这意味着自旋的电场和磁场的贡献相互抵消。第三，空虚空间中场强涨落引起的平方发散的额外能量，由$W_{\mathrm{fl}}=e^2/\pi hcm\ \lim_{p=\infty}p^2$给出。

魏斯科普夫进一步详细分析了狄拉克真空绘景造成的这一新情况。他注意到，首先，泡利不相容原理隐含着，在两个等自旋的电子之间有一个"排斥力"。这个"排斥力"阻止两个粒子靠得太近，相距不能超过大约一个德布罗意波长。作为这种情形的一个结果，人们将会在这个电子的位置上发现真空电子分布中的一个"空穴"，它完全抵消了电子的电荷。但是，同时，人们也会发现在电子的周围有一个来自移位电子的更高的电荷密度云。因此，电子电荷会扩展到h/mc数量级的范围。魏斯科普夫证明，"这一电荷分布的扩展刚好足以把静电自能(从初始的线性发散)约化到对数发散的表达"，$W_{\mathrm{st}}=mc^2(e^2/\pi hc)\log(p+p_0)/mc$[这里$p_0=(m^2c^2+p^2)^{1/2}$]。

其次，不相容原理也隐含着，在初始电子附近发现的真空电子，有一个与初始电子的涨落相位相反的相位涨落。这个相位关系适用于自旋的环形涨落，通过干涉的方式减少它的总电场，但不改变自旋磁场，因为后者由环形电流产生且不依赖于圆周运动的相位。因此，如果给真空增加一个电子，总电场能将通过干涉而减少。也就是说，在量子场论中，一个电子的电场能是负的。魏斯科普夫的精确计算给出$U_{\mathrm{el}}=-U_{\mathrm{mag}}$。与单电子理论中的相应结果($W_{\mathrm{sp}}=U_{\mathrm{el}}-U_{\mathrm{mag}}=0$)相比较，我们发现，在量子场论中自旋对自能的贡献没有消失：$W_{\mathrm{sp}}=U_{\mathrm{el}}-U_{\mathrm{mag}}=-2U_{\mathrm{mag}}=-1/m(e^2/\pi hc)\lim_{p=\infty}[pp_0$ [189]

$-(m^2c^2/2)\log(p+p_0)/mc]$。

再次，量子场论中的能量 W_{fl} 与单电子理论中的相应能量没有什么不同。因此，加上自能的这三部分的最终结果就是这样的。由于 W_{sp} 中的二次发散项被 W_{fl} 所平衡，因此 W_{sp} 的余项之和与 W_{st} 的总值为：$W = W_{st} + W_{sp} + W_{fl} = [(3/2\pi)(e^2/hc)mc^2 \lim_{p=\infty} \log(p + p_0)/mc + \cdots$。而且魏斯科普夫也证明，自能发散在近似值的每一阶上都是对数发散的。

魏斯科普夫的贡献在于：他发现在量子场论中，电子的电磁行为不再完全是类点的，而是在一个有限区域上扩展的。正是这一扩展和随之而来的自旋能量中涨落能量和二次发散项之间的抵消，使自能的发散保持在对数发散的范围之内。这一结果是 20 世纪 40 年代质量重正化理论的一个起点，并且在重正化理论的进一步发展中起着重要作用。值得注意的是，魏斯科普夫的分析完全是建立在狄拉克真空思想之上的。这一事实令人信服地证明以下这一断言是正确的，即狄拉克真空思想确实为量子场论提供了一个本体论基础，尽管一些物理学家包括魏斯科普夫本人，否认狄拉克思想的重要性。

狄拉克真空的无穷极化

然而，狄拉克真空思想一发表，就因它导致了一种新的与电子电荷相联系的发散，而受到了奥本海默的批评(1930b)。奥本海默指出，尽管我们能忽略自由态情形中真空电子分布的可观测效应，但是当外部电磁场出现时，我们就不能避免真空电子引起的无穷大静电场分布的无意义的“可观测效应”。唯一出路似乎是引入“一个无穷大的正电流密度以抵消负电子”(同上)。

这一进展是奥本海默在稍后与弗里关于同一主题的工作中，基于

电子和正电子之间有一种完全的对称而作出的。然而，与电荷相联系的无穷大困难仍然保留，虽然在这种情形中它是另一种困难。在对狄拉克真空中由外部电磁场产生的电子—正电子虚粒子对对于真空中能量和电荷产生的影响进行研究的过程中，奥本海默和弗里 [190] (1934)发现，虚粒子对的能量和外部场的能量之比是$(-\alpha K)$，其中 $K=(2/\pi)\int_0^\infty p^2(1+p^2)^{-5/2}\mathrm{d}p+(4/3\pi)\int_0^\infty p^4(1+p^2)^{-5/2}\mathrm{d}p$。这个有效电荷也将受到因子$(-\alpha K)$的影响。对于大的 p 值来说，K 中的二次积分对数发散。他们认为，这一发散植根于"现有理论一个真正的限制"，并"倾向于在所有涉及极端小长度的问题中出现"(同上)。

然而，关于这种新发散的最初工作，在 1933 年 10 月召开的索尔维会议期间，已由狄拉克本人作出。狄拉克首先引入密度矩阵[通过运用哈特里—福克近似(Hartree-Fock approximation)，分配给每个电子一个本征函数]。然后，他研究了外部静电势对一阶微扰计算中的真空电荷密度的影响，并表明在(产生外部势能的)电荷密度 ρ 和(由真空中的外部势能产生的)电荷密度 $\delta\rho$ 之间的关系是 $\delta\rho=4\pi(e^2/hc)[A\rho+B(h/mc)^2\Delta\rho]$，其中 B 是一个有限常量，A 是对数无穷大。通过在导出 A 的积分中取一个适当的截止(cut-off)，狄拉克发现：

> 通常在电子、质子或其他带电粒子上观测到的电荷，不是实际上为这些粒子携带并出现在基本方程中的真实电荷，而是相差1/137因子的一个稍小的值。
>
> (Dirac，1933)

狄拉克提出的可观测量和出现在基本方程中的参量之间的区分，在重正化思想的发展中无疑是具有重要意义的。然而，在同一论文中出现的对截止的依赖性，表明这一工作并不满足相对论。

在 1934 年的论文中，狄拉克对于真空中电荷密度的特性进行了研究。由于密度明显是无穷大的，因此“现在提出的问题本身是：找到一些消去无穷大量的自然方式……以留下一个有限的余项，这样我们可以假设这个余项为电流密度和电密度”（同上）。这个问题要求他仔细研究光锥附近的密度矩阵中的奇点。结果是密度矩阵 R 可以自然地分成两部分，$R = R_a + R_b$，其中 R_a 包含所有的奇点，并且对于所有的给定场，它也是完全确定的。因此，在电子和正电子的分布中，人们可能做的任何变更将只是对应于 R_b 中而不是 R_a 中的一个变更；只有那些产生于电子和正电子的分布且与 R_b 相对应的电密度和电流密度，才是物理上有意义的。这暗示着一种消去无穷大的方式。[191]

这里，无穷大的消去就意味着只是从 R 中减去R_a。作这种消去的理由是，R_a 与电分布和电流分布中的改变没有任何关系。这样一种减去法的思想尽管总体上不合理，但是在 20 世纪 30、40 年代一直是量子场论中用来处理发散困难的基本方式，直到 20 世纪 40 年代晚期重正化理论取得了重要进展，这一思想才被抛弃。狄拉克所提议的减去法程序为海森伯(1934)所推广；海森伯主张，所有出现在表达式中的发散项都应该被自动地减去，而不需使用截止技巧。这似乎相似于后来的重正化程序，但没有后者的物理基础。

派尔斯对于电荷重正化的更深刻理解

同年，派尔斯(Rudolf Peierls)也发现了一个包含在真空极化表达式中的对数无穷大量，它与自能发散没有直接的联系。关于这一发

散的原因，派尔斯说：

> 发散与这样一个事实相联系，即对电子来说有无穷多态，并且外部场能够引起任何态之间的跃迁。
>
> (1934)

这依次又是与点耦合相联系的，点耦合要求具有很高能量和动量的虚量子承担一种不实在的贡献。派尔斯通过截止在一个动量上的积分（这个动量与等于经典电子半径的波长相对应），发现极化将处在1%的量级上。这一定与某一介电常量相对应，它通过一个常量因子来约化任何“外部”电荷或电流的场。因此，派尔斯声称，我们应该假定，我们正在处理的所有电荷“实际上”比我们观测到的更大，这个假定与狄拉克的假定相一致。

无意义的发散的存在表明，这个理论是不完美的，需要进一步改进。然而，派尔斯说：

> 不知道这一理论中必要的改变，只是一个形式性质的改变，即只是用来避免使用无穷大量的数学上的改变，还是在这些方程下面的基本概念必须进行本质上的修正。
>
> （同上）

这些话语表明，同弗里和奥本海默一样，派尔斯对量子场论抱有一种悲观和怀疑的态度，因为它存在发散困难。我们注意到，有趣的是，同狄拉克、弗里和奥本海默一样，尽管派尔斯要求对量子场论进行激进的革新，但是他提出的获得一个有限结果的方法，并没有改变 [192]

量子场论的基本结构，只是在动量积分中引入一个截止，而这却付出了破坏相对论性的代价。

克拉默斯对于重正化的建议

1936年，人们已表达了这样的猜测，即高动量光子的无穷大贡献与下列量均有联系，它们是无穷大的自质量、无穷大的固有电荷、不可测量的真空量，诸如真空的介电常量。[22]因此，一个系统化的理论，即重正化理论，似乎应该发展为一个防止这些无穷大量出现的理论。沿着这一方向，克拉默斯(1938a)明确提出，必须把无穷大的贡献从真正重要的那些贡献中分离出来并减去，这样，一些可观测效应，诸如"帕斯特纳克效应"(Pasternack effect)[23]，就能够计算(1938a)。他(1938b)又提出，电磁质量，通过消除这一电子的固有场，必须包含在涉及物理电子的理论中。尽管在相对论性量子论场中，这一消除是不可能的，但是，对理论应该具有一个免除"固有场"的特性的认识，对于随后的发展影响重大。例如，它成为施温格(Julian Schwinger)在形成其重正化理论(Schwinger，1983)中的一个指导原则。

散射的研究

在20世纪30年代后期，在散射过程中出现了另一种无穷大特征。[24]首先，布劳恩贝克(Braunbeck)和魏因曼(Weinmann)发现，对于$\alpha(=e^2/hc)$量级散射截面的非相对论性辐射校正，是对数发散的(Braunbeck and Weinmann，1938)。泡利和菲尔兹(Pauli and Markus Fierz，1938)把这个结果看作α的幂展开式不可能正确的证据。相反，他们进行了一个非相对论性的计算，没有涉及α的幂展开式。他们使用了一种后来在重正化过程中广泛使用的接触变换(contact transformation)方法，来分离无穷大的电磁质量并保留了余

项，以使只有散射过程的特性得以表示。这个结果是荒谬的：高频效应将使整个截面消失。泡利和菲尔兹在这一结果中看到了量子场论不适当性的另一个例证。

[193] 为了查明相对论效应的因素将修正泡利和菲尔兹的结论的程度，丹科夫(1939)相对论性地处理了这个问题。他把相对论效应分为三个部分：(i)描述对出现在非相对论性理论中的项进行的相对论性修正的部分；(ii)包括在初始和终了波函数中被散射势能所散射的电子或正电子对部分。(iii)包括为散射势能所散射的产生和湮灭对部分。他的计算表明，在部分(i)和部分(ii)中的每一项都对数发散，但是在每一部分中，项的合并给出了一个有限的结果。在部分(iii)中，情形更复杂：总共有十项。其中六项是有限的，但是由其余四项给出的散射截面的修正，它们的和是对数发散的。

所有与这些修正相关的过程，都涉及负能电子到正能态的跃迁，受由入射电子产生的辐射场的影响。注意到丹科夫在脚注中提到的事情是有趣的：

> 瑟伯(R. Serber)博士已指出，对源于与散射势能场中虚粒子对库仑相互作用的 α 阶散射截面的修正，应该在这里适当地加以考虑。
>
> (Dancoff，1939)

但是，只考虑了由三项组成的这类相互作用的一项之后，丹科夫就得出结论："与这类相互作用相对应的发散项直接源于真空的极化公式，可以被一个适合的电荷密度重正化消除。"(同上)朝永振一郎(Sin-itiro Tomonaga，1966)评论说，如果丹科夫不是那么粗心的话，

那么“重正化理论的历史将会完全不同”。我们将简短地给出这一说法的理由。

补偿对消法

到20世纪30年代末，产生于场的反作用(粒子和它们自己的电磁场之间的相互作用)的各种发散，已使许多物理学家对量子场论失去了信心。在这样的气氛中，一种引入一个“补偿场”(compensative field)的新尝试提了出来。波普(1940)是第一个提出有高阶导数的经典场论的物理学家，在这一经典场论中，出现在达朗贝尔方程(D'Alembert's equation)中的算符□由算符$(1-\Box/k_0^2)\Box$代替。一[194]个满足四阶微分方程的场能被约化为两个满足二阶方程的场。也就是说，新场的势能A_μ能被写作$A_\mu = A'_\mu - A''_\mu$，其中$A''_\mu = (-\Box/k_0^2)A_\mu$。显然，势能$A'_\mu$满足麦克斯韦方程，$A''_\mu$满足一个质量为$kh/c$的矢量介子的方程。因而，新场中点电荷的标势$\phi$为$\phi = \phi' - \phi'' = e/r(1-e^{-kr})$。当$r \gg 1/k$时，$\phi = e/r$，即通常的电磁标势；当$r = 0$时，$\phi = ek$，这意味着$1/k$充当电子的有效半径。如果矢量介子场提供一个负“机械”压力(它对于补偿库仑排斥力是必要的)，那么这个点电子在新场的作用下将是稳定的。在这种情形中，电子的自能是一个有限量。

波普的理论满足相对论，并能推广到量子情形。在这个理论中，纵场的自能发散能被抵消，但这个理论仍然不能处理横场自能的发散，这是量子场论中的显著困难。尽管有这一困难，波普从庞加莱那里继承来的补偿思想，对于派斯(1945)、坂田(1947)、朝永(1948； with Koba，1947，1948)和其他人有直接且巨大的影响。而且，费恩曼保持相对论不变性的截止方法(1948a，b)和泡利—维拉斯(Villars)的规则化方案(1949)，在为了消除奇点而引入一个辅助场的

意义上，与波普的理论等价。

受波普的直接影响，派斯(1945)发表了一个克服自能发散困难的新方案。这个新方案把每个基本粒子看作一组场的源；用这种方式，各种对这些场引起的自能无穷大作出贡献的部分相互抵消，从而使得最终的结果变得有限。关于电子的具体情形，派斯假定电子除了是电磁场的源之外，还是一个中性的、短程的矢量场的源。这个"减去的"矢量场直接与电子耦合。如果收敛关系 $e = f$（e 是电荷，f 是电子与矢量场耦合中电子的电荷）得到满足，那么电子的自能对任何阶近似都是有限的。然而，派斯承认这样一种减法场与空穴理论要求的真空电子的稳定分布不相容。因此，在量子场论的框架中这个方案不得不舍弃。

坂田(1947)独立地获得了派斯已获得的相同结果。他命名辅助场为内聚力场或C介子场。然而，坂田对自能发散的解答实际上是一个幻想。正如木下东一郎(Toichiro Kinoshita，1950)表明的，当讨论扩展至最低阶近似之外时，两个耦合常数之间的必要关系式将不再抵消发散。不过，把C介子假说用作一个催化剂却是富有成效的，[195] 它导致朝永引入了自洽的减去法。

受坂田用C介子场作为一个补偿场来"解决"自能发散所获得的成就的启示，朝永试图把坂田的思想扩展到散射问题。最初他们没有成功(Tomonaga and Koba，1947)，因为朝永和木庭(Koba)在他们的计算中重复了丹科夫遗漏了某些项的错误。但是，不久之后，通过运用一种新的有效得多的计算方法(见下文)，他们发现，比较在丹科夫的计算和他们自己以前的计算中出现的各项时，有两项被忽略了。虽然只有两项被遗漏，但是它们对于最终结果却是至关重要的。在他们纠正了这一错误之后，在电子的散射过程中出现的无穷

大，由于电磁场和内聚力场而完全抵消了，除了真空极化类型的发散之外(这类发散能通过重新定义电荷而立即被消除)。这样，朝永通过结合两个不同的思想——电荷的重正化和C介子场的补偿机制——对电子的散射实现了一个有穷校正。

看来在这里作一下评论是适宜的。补偿对消思想的提出者和追随者的关注点，与我们将要简短讨论的重正化思想的关注点不同，他们不是分析和仔细应用有关耦合电子和电磁场的已知相对论性理论，而是改变它。他们引入未知粒子的场，以这样一种方式取消由已知相互作用产生的发散，因此超越了现存的理论。因此，对量子场论来说，这是一种与重正化方法(它是在现存理论中所用的一种策略)不同的方法，虽然这两种方法由于它们消除发散的目标相似而容易彼此混淆。

质量重正化的贝特—刘易斯方案

成熟的重正化思想始于贝特(Hans Bethe，1947)和刘易斯(H. A. Lewis，1948)的工作。在兰姆和雷瑟福德(Lamb and Retherford，1947)用微波方法重新发现氢的能级移位之后不久，贝特接受了施温格和魏斯科普夫的建议，尝试着用辐射场解释由电子相互作用造成的移位。这一移位在所有现存理论中都出现无穷大，因此不得不忽略掉。然而，贝特指出：

> 用一个电磁质量效应来辨别能级移位中最强的发散项(在非相对论性计算中是线性发散，在相对论性计算中是对数发散)，是可能的。我们应恰当地认为，这种效应已经包括在电子的观测质量之中，因此它必须从理论表达中消去。这样，在非相对论性理论中结果只是对数发散而不是线性发散，而在其最强的发

[196]

散只是对数无穷大的相对论性理论中收敛。这将对光的频率设置了一个数量级为 mc 的有效上限，它有效地贡献出一个束缚电子的能级移位。

(Bethe, 1947)

在解释了存在这样一个上限之后，贝特运用这个限制截止了在非相对论性计算中对数发散的积分，获得了一个与实验值(1040 兆周)极好地符合的有穷结果。

贝特是把电磁质量效应从量子过程中明确区分开来的第一人。他把电磁质量效应合并到源于观测质量的效应之中，而不是简单地忽略它。除了电子质量本身的重正化外（这一点或多或少地为他的先辈所讨论过），贝特的质量重正化程序也能应用于各种量子过程。以前，只是因为无穷大量不可观测，就不讲道理地减去它们，这就会使产生无穷大量的这个理论是否有效的问题悬而未决。与上述做法不同，他们的“合并”程序已为对于发散的进一步物理解释铺平了道路。

受贝特的成功的鼓舞，刘易斯(1948)把质量重正化程序应用到对散射截面辐射校正的相对论性计算中。在这样做的过程中，刘易斯使重正化思想的本质变得更为清楚。他主张，隐藏在整个重正化纲领之下的关键假设是：

电子的电磁质量是一个微小的效应，它表观上的发散源于当前量子电动力学超过某个频率时的失效。

(Lewis, 1948)

显然，只有当一个物理参量（在量子场论的微扰计算中，这个参量可证明是发散的）确实是有限和微小时，它与“裸”参量的分离和合并才能被看作数学上合理的。量子场论在极端相对论性能量（ultra-relativistic energies）处的失效，正如微扰理论中发散所表明的，隐含着现存量子场论框架有效的区域应该与它失效的区域区分开来，在现存量子场论失效的区域，新的物理学将显示出来。虽然不可能确定界线在哪里，人们不知道什么理论可以用来计算在量子场论中不可计算的小效应。[197] 然而，可知和不可知的分离可以通过引入一个截止点而在数学上加以实现，可以通过使用唯象的参量而图式化，在这个图式化中，必定包括这些小的效应。

这样一个重正化微扰理论（其中量子场论的表述与未加解释的观测质量值相混合），有时被称作不完备的理论，或合理的工作假说。但是，它在理论上比以前的所有理论，诸如派斯或坂田的更为合理；在那些理论中，两个有相反符号的无穷大量相互结合，以给出一个有限的量。那些理论只能被看作为了解决发散问题而作的一种早期智力探索，但在物理上和数学上是不能令人信服的。与那些理论相反，刘易斯把假设为很小的电磁质量与有限的“力学”质量相结合，以获得观测质量，同时从相互作用中减去电磁质量的效应。用这种方式，重正化程序获得了坚固的物理基础。因此，上述提议应该被看作在重正化纲领形成过程中的一个重要事件。

刘易斯还证明，贝特关于在量子过程的计算中电磁质量产生发散项的猜测是正确的。刘易斯在相对论性计算中，发现电磁质量确实导致一个只在倍增常数方面与丹科夫发现的发散辐射校正项不同的表达式。唯一的数值差异来自为丹科夫所忽略的某种静电跃迁，一旦再次考虑被丹科夫忽略的项，这一数值差异将消失。因此，刘易斯

相当坚定地认为，由丹科夫发现的那些发散项，可以看作与电子电磁质量的表现形式是同一的。因此，如果有人在哈密顿量中使用了经验质量，包括电磁质量，那么他一定是忽略了源于电磁质量效应的跃迁。总之，使用经验质量并忽略源于电磁质量的效应，是贝特—刘易斯质量重正化程序的本质内容。

正则变换

在贝特—刘易斯的方案中，关键点是一个形式上无穷大的项的有穷部分的无歧义分离。刘易斯承认，由施温格发展的关于哈密顿量正则变换(canonical transformation)的方法，正好适合这一目标。施温格(1948a)提出，刘易斯提到的分离能通过如下方式实现：

> 变换当前流行的空穴理论电动力学的哈密顿量，以清楚地展示自由电子源于光量子的虚发射和虚吸收而对数发散的自能。 [198]
>
> (同上)

施温格认为，新哈密顿量在以下三个重要的方面优于原初的哈密顿量：

> 它涉及实验电子质量，而不是不可观测的力学质量；电子如今只在外部场存在的情况下与辐射场相互作用……；电子与外部场的相互作用能量，如今受制于有限的辐射校正。
>
> (同上)

对于由真空极化产生的对数发散项，施温格和他的前辈一样，认为：

这样一项相当于用一个常量因子改变电荷的值，只有终值与实验电荷完全一致。[因此]所有的发散都包含在重正化因子之中。[25]

（同上）

直到 1948 年 3 月 30 日到 4 月 1 日波科诺会议召开，施温格才发表他的正则变换。尽管朝永在 1943 年到 1946 年期间已发展了相同的方法(见下文)，但由于那时的通讯条件，这一方法在美国不可能获悉。运用这一方法，施温格(1948b；1949a，b)从电磁场和正电子—电子场之间的对称性观点出发，系统地考察了量子场论中的发散。根据他的研究，出现发散的基本现象是真空极化和电子自能。这两种现象非常相似，本质上描述了每个场与另一个场的真空涨落的相互作用：由电磁场产生的正电子—电子场的真空涨落，是电子—正电子对的虚产生和虚湮灭；由电子产生的电磁场的真空涨落，是光子的虚发射和虚吸收。施温格总结说：

即使依据对数发散因子，这些涨落相互作用的效应也仅仅是改变基本常量 e 和 m，当前理论的所有物理上重要的发散都包含在电荷和质量重正化因子之中。

（同上）

以对各种发散现象的这样一种完整和彻底的理解为基础，并配备[199] 先进的数学工具，诸如正则变换、变分原理和函数积分，施温格得以成功地处理一系列困难问题，诸如能级移位和反常磁矩。他的结果与实验符合得很好。这一事实已强有力地表明了重正化微扰理论的

巨大威力，也激励了物理学家对重正化做进一步的彻底研究。

朝永的贡献

尽管朝永只是在贝特和刘易斯的工作发表之后，并且是在他逐渐放弃了补偿场的天真思想之后，才对重正化获得了一个真正的理解，但是他在对重正化纲领的形成而言非常重要的一些技术发展中，的确与贝特和刘易斯享有同等的和独立的地位。早在 1943 年，朝永就用日文发表了一篇关于量子场论的相对论性不变的论文[26]，在其中所做的相对论性推广中，利用了诸如正则对易关系和薛定谔方程之类的非相对论性概念。通过这样做，他把一个幺正变换(unitary transformation)应用到所有场，为这些场提供非相互作用场的运动方程；而在这个变换的(态函数的)薛定谔方程中，只有相互作用项保留了下来。也就是说，他利用了相互作用表象(interaction representation)。然后，他注意到，当考虑一个非相互作用场时，对于任意具有四维特征的时空点，在展示对易关系上不存在困难。至于变换的薛定谔方程，他用“一个类空表面”取代时间，使它获得相对论协变性。

然而，在处理发散时，朝永没有利用他强有力的技术方法。相反，他只是诉诸补偿的思想，直到贝特的工作发表，才强烈吸引了他的注意：

> 因为它也许表明了一条克服量子场论基本困难的可能路径，并且第一次与可靠的实验数据发生了密切的联系。
>
> (Tomonaga and Koba，1948)

于是，朝永提出了一个形式体系，他称之为“自洽的减去法”，以一种更精密、更合乎情理的形式来表达贝特的基本假设；在这个方法

中，被减项之间的分离由一个正则变换来完成。然后这个形式体系被应用到散射问题，以隔离电磁质量项，并借助于C介子场而获得一个有限的结果（Ito, Koba, and Tomonaga, 1948）。在朝永和木庭的论文（1948）中，这个C介子理论仍被看作旨在实现减去法运算，以获得一个有限的自能和有限的散射截面。但问题在于这似乎是一个特设的策略。如果目标只是处理量子场论中的发散困难，那么C介子场的引入远不如重正化程序方便；更糟糕的是，它具有不可能克服的由真空极化产生的发散困难。因此，C介子理论或者更普遍的补偿思想，逐渐在朝永和其他日本物理学家的工作中被重正化程序所取代。

[200]

费恩曼的调整

上面我们提到，重正化程序的一个必要假设是：在量子场论计算中明显的和不适当的发散量，实际上是有限的小量。只有在这种情形中，人们才能把这些发散量从量子场论获得的表达式中合理地分离开来，而把它们融合到裸量中，并用观测值取代总和。因此，为了能使发散量的分离和合并在数学上变得合理，不得不建立一组调整规则，使得有可能以一种保持相对论性不变性和规范不变性的方式来计算物理量。这样一组调整规则是由费恩曼基于他对于相应的经典情形的讨论（1948a）而提出的（1948b）。

费恩曼把这组规则叫做相对论性截止（relativistic cutoff）。这一技巧用一个有限的截止，把本质上纯形式的发散量的操作，即参量的重新定义，变换成为相对不错的数学运算。这一思想的要点如下：对于一个辐射场的真空涨落，我们用新的密度函数 $g(w^2 - k^2) = \int_0^\infty [\delta(w^2 - k^2) - \delta(w^2 - k^2 - \lambda^2)] G(\lambda)\, d\lambda$ 替换旧的密度函数 $\delta(w^2 - k^2)$。这里，$G(\lambda)$ 是一个光滑函数，且 $\int_0^\infty G(\lambda)\, d\lambda = 1$。

在光子传播子的动量空间表象中，这意味着用 $\int_0^\infty [1/k^2 - 1/(k^2 - \lambda^2)]\, G(\lambda)\, d\lambda$ 替换 $1/k^2$。在数学上，这等于用收敛因子 $c(k^2) = \int_0^\infty - \lambda^2(k^2 - \lambda^2)^{-1}\, G(\lambda)\, d\lambda$ 乘 $1/k^2$。

对于电子场的真空涨落，费恩曼试图用相似的方法，即对描述真空极化的表达式 $J_{\mu\nu} = - \mathrm{i}e^2/\pi\int_0^\infty \mathrm{sp}[(\gamma \cdot p + \gamma \cdot q - m)^{-1}\gamma_\mu(\gamma \cdot p - m)^{-1}\gamma_\nu]\, d^4 p$ 中电子的两条内线，分别引入收敛因子 $c(p^2 - m^2)$ 和 $c[(p + q)^2 - m^2]$。这里，sp 意思是迹(spur)，$(\gamma \cdot p + \gamma \cdot q - m)^{-1}$ 和 $(\gamma \cdot p - m)^{-1}$ 是动量分别为 $(p+q)$ 和 p 的电子的内线动量空间表象(Feynman, 1949b)。这种方法只在以破坏流守恒和规范不变性为代价时才使得 $J_{\mu\nu}$ 收敛。因此，必须放弃这种方法。然而，费恩曼意识到，收敛因子的引入在物理上隐含着，一个质量为 m 的粒子的内线贡献将被一个质量为 $(m^2 - \lambda^2)^{1/2}$ 的新粒子的贡献所抵消。考虑到真空极化图中任何内线将形成一个闭环这个事实，我们不应为两条内线引入不同的收敛因子。相反，我们应该为闭环引入一个质量为 $(m^2 - \lambda^2)^{1/2}$ 的新粒子的贡献，以获得一个收敛的结果。把 $J_{\mu\nu}$ 写作 $J_{\mu\nu}(m^2)$，并用 $J^p_{\mu\nu} = \int_0^\infty [J_{\mu\nu}(m^2) - J_{\mu\nu}(m^2 + \lambda^2)]\, G(\lambda)\, d\lambda$ 代替它，计算结果除了一些项依赖于 λ 之外，一定收敛。 [201]

对于 λ 的依赖似乎破坏了相对论性，正如以前引入的任何截止所做的一样。而且，在 λ 趋于无穷大的极限时，这一计算仍然导致发散表达式。事实上，情况并非如此。明显依赖于 λ 的这个量，即使它是发散的，也能在重正化后被吸收，使以这种方式引入的截止与相对论性不冲突。特别地，如果在质量和电荷重新定义之后，其他过程对于截止值不敏感，那么重正化理论就可以通过让这个截止趋于无

穷大而得到确定。如果一组有限数量的参量足以确定一个理论为一个重正化理论，那么这个理论就叫做可重正化的。

从物理上讲，费恩曼的相对论性截止等同于引入一个辅助场(及其相关粒子)，以抵消由于原初场的("实")粒子产生的无穷大贡献。[27] 费恩曼的方法不同于实在论的调整理论或补偿理论。在补偿理论中，具有有限质量和正能的辅助粒子，被假定为原则上是不可观测的，并且由明确进入哈密顿量的场算符描述。在如下意义上，费恩曼关于截止的理论是形式化的：辅助质量仅仅被用作数学上的参量，这些数学参量最终趋于无穷大，并在原则上是不可观测的。"实在论"方法以下面这些作者的论文为代表：坂田(1947，1950)、梅泽(Umezawa, Yukawa and Yamada, 1948; Umezawa and Kawabe, 1949a, b)和其他日本物理学家，以及拉伊斯基(Rayski, 1948)。我们发现在"形式论者"(formalist)之中，除了费恩曼外，还有里维耶和斯蒂克尔伯格(Rivier and Stueckelberg, 1948)，泡利和维拉斯(1949)。

费恩曼对于重正化纲领的其他贡献与物理过程计算的简化相关。他不但为因果传播子提供了简洁的形式，而且也提供了一组图解规则 [202] (diagram rules)，以使在描述物理过程的 S 矩阵元中的每个因子，以一对一的方式与平面上的线或顶点相对应(1949b)。图解规则是一个方便、有力的工具，使得费恩曼能够表达和分析用量子场论描述的各种过程，并且清楚地体现了朝永、贝特、刘易斯和施温格关于正则变换、发散的分离和重正化的思想。所有这一切为戴森(Freeman Dyson)通过进一步分析和组合图解的贡献而提出一个自洽且完备的重正化纲领提供了先决条件。

戴森的重正化纲领

戴森综合了贝特、刘易斯、朝永、施温格和费恩曼有关重正化的

思想和技巧。他利用费恩曼图，详细分析了出现在S矩阵元计算中的各种发散；在他1949年的论文中系统地提出了处理这些发散的重正化纲领(1949a)，并且在同年的另一篇文章中对这个纲领做了一些补充和完善(1949b)。

戴森的纲领完全基于对费恩曼图的分析。首先，他论证了不连续的曲线图的贡献可以忽略掉。其次，他定义了初始发散曲线图，所有发散曲线图都可以约化成该初始发散曲线图。于是，通过分析这一图式的拓扑，他获得了那些发散曲线图的收敛条件 $k = 3E_e/2 + E_p - 4 \leqslant 1$，这里 E_e 和 E_p 分别是图式中的外电子线数和光子线数。运用这一收敛条件，戴森把所有可能的初始发散分成三类，即电子自能、真空极化和电磁场中单电子的散射。在进一步的分析中，戴森引入了关于图式的构架的思想，可约化和不可约化的曲线图、规范和不规范的曲线图，等等。运用这些思想，他提出在计算一个S矩阵元时，我们应该(i)画出相关的不可约化曲线图；(ii)将观测质量代入哈密顿量中；(iii)用新的传播子 $S'_f = Z_2 S_f$，$D'_f = Z_3 D_f$ 和新的顶点 $\Gamma'_\mu = Z_1^{-1} \Gamma_\mu$ 取代最低阶传播子S_f、D_f 和 Γ_μ(它们不会导致发散)[28]；(iv)用新的裸波函数 $\Psi' = Z_2^{1/2} \Psi$，$\Psi'^- = Z_2^{1/2} \Psi^-$ 和 $A'_\mu = Z_3^{1/2} A_\mu$，代替不会引起发散的 Ψ、Ψ^- 和 A_μ，其中 $Z_{1,2,3}$是"发散因子"。戴森声称，这样做了以后，所有产生于两类真空涨落的三类初始发散，即所有的辐射修正，就都得加以考虑，并会在质量重正化之后导致一个"发散因子"$(Z_1^{-1} Z_2 Z_3^{1/2})^n$。然后考虑在第 n 阶过程中，必定有一个因子 e_0^n 出现(其中 e_0 是出现在更为基本的方程中的所谓裸电荷)，并假设观测电荷 $e = (Z_1^{-1} Z_2 Z_3^{1/2}) e_0$，那么将不再出现任何发散困难。 [203]

显然，这个程序在精细结构常量 $\alpha = e^2/hc$ 之内，可以被无限制地用到任何阶。人们不禁会问：为什么戴森称这样一个优美的方案为一个纲领而不是一个理论呢？他的回答是他没有给出有关高阶辐射修正的收敛的普遍证明。也就是说，虽然他的程序保证了在一个任意阶近似中获得的辐射修正是有限的，但是它不能保证对这些辐射修正的求和仍然是有限的。

总结

重正化纲领的形成由三个主要步骤组成。第一，揭示各个发散项的对数性质。第二，把所有发散项约化为两个源于自能和真空极化的初始类型。第三，通过(i)应用正则变换以得到量子场论的协变形式(相互作用表象)并且分离发散项；(ii)质量和电荷重正化；(iii)连贯运用质量和电荷的经验值；从而找到处理发散量的一个无歧义的、连贯的方式。借助于这一重正化纲领，可以获得有限辐射修正，并且发散困难问题也能得到处理。因此，量子场论成为预测各种物理过程的强有力的计算手段，也成为当代物理学中最具预测力的学科分支之一。

狄拉克对重正化的批评

具有讽刺意味的是，重正化思想的积极开拓者之一狄拉克(1963，1968，1983)，不顾重正化思想取得的所有这些巨大成功，反而对此进行了猛烈攻击。在狄拉克看来，所有这些运用重正化获得的成功，既没有一个可靠的数学基础，也没有一个令人信服的物理图景。从数学上看，重正化要求忽略无穷大量而不是无穷小量，这与数学中的典型习惯相反。因此，这是一种人为的或不合逻辑的做法(Dirac，1968)。从物理上讲，狄拉克认为无穷大量的存在暗示着：

> 我们关于电磁场与电子的相互作用理论存在某些根本性错误。鉴于这些根本性错误，我的意思是说，这一机制是错误的，或者说这一相互作用机制是错误的。
>
> (Dirac, 1983)

然而，重正化程序与对理论中的错误的修正没有关系，或者说与搜索新的相对论性方程和新类型的相互作用没有关系。相反，它只是一个给出实验结果的笨拙的处理规则。因此，它将误导物理学走上一条错误的道路(Dirac, 1963)。狄拉克强烈谴责重正化并强调说："这在物理上是彻头彻尾的胡说，人们应该做好彻底放弃它的准备。"(Dirac, 1983) [204]

朝永的态度没有这样消极。但是他也承认：

> 诚然，我们的(重正化)方法绝没有真正解决量子电动力学的基本困难，但是在没有触及基本困难的前提下，它给出了处理场效应问题的一种没有歧义的、连贯的方式。
>
> (Tomonaga, 1966)

为重正化纲领辩护

固然，在狄拉克对重正化的批评中有一些深刻的见解。但总体看来，他的批评似乎有失公允，因为他的批评忽略了重正化的概念基础。事实上，重正化纲领能够在两个层面上得到辩护。

从物理上讲，高阶无穷大量全都与高动量虚光子(和电子—正电子对)的无穷大贡献相关联，表明量子场论的形式体系包含来自与高动量虚光子(和电子—正电子对)相互作用的非实在贡献。虽然虚量

子过程的物理实在性为实验所证实[29]，但是无穷大动量虚量子显然是非实在的。不幸的是，这个非实在的要素深深植根于量子场论的基础之中，植根于算符场和由定域算符场所产生的局域激发的概念之中。各种格林函数是这种局域激发之间的相关函数，而研究它们的时空行为，对于识别这些物理粒子和它们之间的相互作用，是唯一的手段。在这种语境下，重正化可以恰当地理解为，把注意力从有关初始局域激发和相互作用的假设世界，转移到物理粒子的可观测世界这样一种视角上来。它的主要目的，正如戴森所说的：

> 不是对当前理论作一彻底修正，将所有无穷大量变得有穷，而是对这个理论作一转向，旨在使有限量成为主要的。
>
> (1949b)

[205] 从哲学上讲，重正化能被看作一个基于原子论(atomism)的上层建构，原子论深深植根于西方的思维模式中。在量子场论框架中发展的各种描述亚原子世界的模型，实际上都保留了原子论的特征：用出现在拉格朗日函数中的场来描述的粒子被看作世界的基本构成。然而，未重正化理论所采纳的原子论或原子性(atomicity)的观念，具有不彻底的特征。这一点已由施温格(1973)清楚地指出，其原因是未重正化算符场理论包含着关于物理粒子的内部结构的隐含推测，即认为物理粒子的内部结构对高能状态的动力学过程的细节是敏感的。从数学上讲，这个假定已表明它自身处于发散积分状态。[30]从形而上学的角度讲，这个假定暗示，用出现在拉格朗日函数中的场来描述的物理粒子，存在更基本的构成，这与粒子或场作为世界的基本砖块的情形相矛盾。

在这方面，重正化程序的基本重要性可用以下方式来描述。通过排除所涉及的任何不可达的、极高的能量区和相关的结构假设，重正化程序强化了量子场论的原子论承诺。这是因为出现在重正化理论中的粒子或场，确实充当着世界的基本砖块。但是注意，这里原子性不再指称严格的点模型。在某种程度上，人们排除了所涉及的不可达的高能区，因为，由于不确定性原理，高能区从概念上讲源自严格的点模型，这样，重正化就模糊了任何类点(point-like)特征。

这种已为重正化量子场论所采纳或产生的空间上延展但又无结构的准点模型(quasi-point model)，可以在两个方面得到辩护。一方面，它有其经验上的成功的支持。另一方面，通过论证，它也能从哲学上得到辩护：只要实验能量不是高得足以达到探测粒子的内部结构，因而关于它们的小距离结构的所有陈述根本就是猜测，那么准点模型不但是一个适合于实验目的的有效近似，而且也是我们必须经历的一个必要的认知阶段。

与狄拉克(1983)和朝永(1966)曾渴望获得的真实理论相比较，并与超弦理论学家仍然在孜孜以求的“万有理论”(theory of everything)相比较，准点模型似乎只是一个在过渡时期需要的数学设计。一位狄拉克追随者会争辩说，一旦基本粒子的结构已知，准点模型就应该放弃。这是真的。但是，那些忠实的原子论者会争辩说，在(目前已被公式化和实践的)物理学中，对任何层次上的物体的结构分析，总是建立在下一层次(似乎)没有结构的物体(真正的点或准点、类点)基础之上。例如，如果没有对类点部分子的弹性散射分析作为基础，那么对于深层的非弹性轻子—强子散射的分析将是不可能的。因此，在量子场论中采纳准点模型似乎是不可避免的。这可能就是费恩曼(1973)作如下陈述时的意思：“我们的出发点是先假设 [206]

[准点粒子]的存在，因为否则我们根本不会有场论。”

在这种语境中，原子论的辩证法可总结如下：一方面，就我们所知，在任何层次上，粒子的无结构特征不是绝对的，而是有条件的和依赖语境的，只被相对低能量的实验探测证明是正确的。当实验中可获得的能量变得足够高时，粒子的一些内部结构迟早会被揭示，绝对不可分的观念将证明是一个幻想。另一方面，随着粒子结构在一个层次上的揭示，同时，作为这种揭示的一个前提条件，在下一个层次上出现了(似乎)没有结构的物体。因此，只要量子场论保留这一表示模式，“结构物体用(似乎)没有结构的物体表达”的原初图式就会保留，并将继续保留。[31]由于如果不重正化，点模型场论就不会有任何意义，因此，狄拉克关于重正化“是在未来不会保持的东西”(1963)的预言，听起来是没有根据的。相反，根据原子论的拥护者，重正化将始终起作用，与定域场论一同起作用。

恼人的问题

另一方面，彻底的原子论需要一种无限的递归，这在哲学上是不那么吸引人的。为了避免这个令人不愉快的结果，20 世纪 60 年代初，人们提出了另一个不同的纲领：S 矩阵理论[S-matrix theory，基于靴袢(bootstrap)思想，它是当代粒子物理学中整体论的一个变种]。[32]靴袢哲学在场论框架以外运行得很好，甚至试图成为一个反场论纲领。S 矩阵纲领为弦理论(string theories)和超弦理论(superstring theories)双重模型(dual model)提供了基础，在这一基础上，弦理论和把超对称性合并进来的超弦理论得以发展。[33]超弦理论的显著特征之一是，其中一些超弦理论已预示有一个不受无穷大阻碍的计算方案。如果这证明是正确的，超弦理论的框架就能发展为一个连贯的框架，那么狄拉克对于重正化的批评将最终证明是正确的和

有洞察力的。

然而，其他令人困扰的问题仍然遗留了下来。首先，众所周知，重正化程序实际上只在微扰框架中起作用。它附随于 e^2/hc 小量(大约为 1/137)。[34]但是，烦恼的是在一个极短程上的有效耦合常数变得比 1 还大，在这种情形中不能使用微扰方法。是否存在一个能被非微扰方法重正化的理论？或者说，由于临界距离小于电子的施瓦氏半径，电动力学和广义相对论未来的统一将会医治发散的痼疾吗？或者说，电动力学和强相互作用的统一将产生这个问题的一个解答吗？无论如何，电磁力和弱相互作用的统一未能用于这个目标，因为它保留在微扰重正化的框架之内。关于这个主题，在 8.2 节、8.3 节、8.8 节和 11.2 节中将会有更多的讨论。 [207]

其次，一个深刻得多的问题与规范不变性有关。量子电动力学在其重正化中是成功的，主要是因为它的规范不变性，它要求其规范量子（光子）质量为零。而有限质量的玻色子一旦进入理论以描述其他基本相互作用，规范不变性就会遭破坏，并且这些理论通常不可重正化。重正化纲领的进一步发展与规范理论密切相关，这将在第三篇仔细研究。

注释

1. 关于本体论的更详细的讨论，见 1.4 节。
2. 见 Rosenfeld(1973)和 van der Waerden(1976)。
3. 相同的说明见 7.3 节。
4. 见 Wentzel(1960)和 Weisskopf(1980)。
5. 这里提到的思想归功于约当，见 van der Waerden(1967)和约当的回忆录(1973)。
6. 在狄拉克的论文(1927b, c)中，"h" 实际上等于 $h/2\pi$。我遵照狄拉克在其讨论(1927b, c)中的用法。
7. 早在 1926 年 8 月，狄拉克已指出："似乎可能建立一个电磁理论，在这个理论中，在时空中的一个特定点 X_0，Y_0，Z_0，t_0 上，场的势能由 X_0，Y_0，Z_0，t_0 的函数——常量元的矩阵——表示。"(1926b)

8. 在玻色子的情形中，量子化条件是 $[\theta_r, N_s] = i\hbar\delta_{rs}$，因此 $b_r b_s^* - b_s^* b_r = \delta_{rs}$。这与具有对称波函数的位形空间波动力学相一致。在费米子情形中，量子化条件必须以 $\theta_r N_s + N_s\theta_r = i\hbar\delta_{rs}$，因此 $b_r b_s^* + b_s^* b_r = \delta_{rs}$ 这样一种方式加以修正。这与具有反对称波函数的位形空间波动力学相一致，在此，泡利不相容原理得到满足(见 Jordan and Wigner，1928)。

9. 正如我上面已指出的，这些关系本身不能被看作量子化条件。

10. 场的零点涨落与零点能没有直接的联系，那纯粹是形式上的特征，能通过被称作"正规编序"(normal ordering)的重新定义而去除(见 Heitler，1936)。

[208] 11. 关于狄拉克对于基质的思想，见他《量子力学原理》(*Principles of Quantum Mechanics*，1930*b*)一书的前言；关于他后来对于非机械以太的相信，见他的论文(1951，1952，1973a)。

12. 例如，杨振宁坚持这一立场。对于狄拉克充盈真空的概念被奥本海默和弗里、泡利和魏斯科普夫的工作之后的进展所拒斥的说法，他对我评论说："这是肤浅的，并没有领会狄拉克思想的精神实质。"(1985 年 5 月 14 日的私人谈话)另见 Lee and Wick (1974)。

13. 根据罗森菲尔德(1968)，这个思想在赫兹的思想中也有其经典渊源。他强调说，赫兹的"中项"(middle term)思想能被解释为"另一种隐粒子"，相互作用经由它得以传递。

14. 从相对论性经典拉格朗日场论出发，海森伯和泡利(1929，1930)发展了一个普适的量子化场论，借助于正则量子化，从而为费米和汤川的相互作用理论准备了工具。

15. 根据维格纳规则(Wigner and Witmer，1928)，一个由奇数个自旋为 1/2 的粒子组成的复合系统必须服从费米统计，一个由偶数个自旋为 1/2 的粒子组成的复合系统必须服从玻色统计。

16. 这一思想的新颖之处得到了多尔夫曼(Yakov Dorfman，1930)的强调。

17. 见 Yang and Mills(1954a，b)。

18. 在查德威克发现中子之后，泡利的"中子"被费米和其他人重新命名为"中微子"。

19. 这个模型被汤川秀澍和坂田昌一(1935a，b)采纳。

20. 见 Hayakawa(1980)。

21. 所谓"二介子假说"(two-meson hypothesis)的提出是为了取代汤川的最初方案：自然界中存在两种质量不同的介子，重介子 π 负责核力，轻介子 μ 是重介子的衰变产物，并在与电子和中微子发生弱相互作用时被观测到。见 Sakata and Inoue(1943)，Tanikawa (1943)，以及 Marshak and Bethe(1947)。

22. 见 Euler(1936)。

23. 帕斯特纳克(Simon Pasternack，1938)讨论了在观测到的氢的两个态 $2S_{1/2}$ 和 $2P_{1/2}$ 之间的能级分裂。这一效应再次为兰姆和雷瑟福(1947)证实，从那以后以"兰姆移位"著称。

24. 在那时，物理学家特别关心散射中的"红外灾难"。但是，这与重正化的讨论不相关，因此这个主题将不在下面的讨论中涉及。

25. 应该提到，施温格可能是充分认识到电荷重正化只是电磁场的一个属性、并导致电荷的分数缩减的第一个物理学家。例如，作为对照，见佩斯给朝永的信，佩斯在信中写道："最令人困惑的问题似乎是如何'重正化'电子和质子的电荷，通过这种方式以使这些量的实验值相互相等。"(1948 年 4 月 13 日，转引自 Schwinger，1983)

26. 英译本发表于(1946)。

27. 应该强调，这里补偿思想的出现纯粹是形式上的，因为正如我们不久将看到的，对 λ 的依赖是暂时的，在最终的表达式中将消失。

28. 这个程序只有当发散子图不重叠时才起作用，而在处理重叠发散时就失效了，这种重叠发散在量子电动力学中只出现在固有自能图中。为了隔离发散图，戴森定义了一个数学程序(没有发表)。萨拉姆(Abdus Salam，1951a，b)把戴森的程序扩展成普遍的规则，并把这些规则应用到标量电动力学的重正化。沃德(John Ward，1951)借助于沃德恒

等式(1950)，解决了电子自能函数中的重叠发散困难。沃德恒等式是规范不变性的一个结果，并且在电子自能函数的导数和若没有重叠问题就能直接重正化的顶点函数之间，提供了一个简单的关系。米尔斯和杨振宁(1966)推广了沃德的方法并解决了光子自能函数中的重叠发散困难。我感激杨振宁博士让我注意这个微妙的主题。 [209]

29. 例如，一个受激氧核的电子—正电子发射对虚光子过程的物理实在性的检验。

30. 赫西(1961)断言，发散困难的根源可以追溯到基本的原子性概念。如果原子性概念指称点模型或局域激发和局域耦合，那么这一断言是正确的。然而，如果如通常所做的，它指称没有结构的基本成分，那么发散困难的根源只是心照不宣地被未重正化算符场论采纳的结构假设，而不是原子性概念。

31. 因此，在万物是可分的S矩阵理论中由靴袢假说表达的思想，与量子场论所采用的原子论范例相矛盾。

32. Chew and Frautschi(1961a，b，c). 另见8.5节。

33. 关于这一评论，见Jacob(1974)；Scherk(1975)和Green(1985)。

34. 因此，严格地讲，汤川的介子理论由于不能重正化(由于强相互作用中的耦合常数很大)，从来就不是一个关于核力的场论——虽然它是在朝着正确方向迈出的一步——因此人们无法根据这一介子理论计算出什么。

第 8 章
量子场纲领

[210] 在量子场论框架中，关于荷电粒子和电磁场之间的相互作用的研究，被称作量子电动力学(QED)。量子电动力学，特别是其重正化的微扰形式，被各种描述其他相互作用的理论所仿效，因而成为一个新的研究纲领——量子场纲领(QFP)——的起点。量子场纲领已由一系列理论贯彻实施，其发展受到它从量子电动力学中继承而来的一些固有特征的强烈限制。由于这个原因，我将从概述这些特征开始，来探讨量子场纲领错综复杂的演变。

8.1 本质特征

量子电动力学是一个由服从运动方程的定域场算符、某些正则对易和反对易关系(分别对应于玻色子和费米子)以及通过把场算符连续地应用于真空态——作为一个缺乏任何物理属性的洛伦兹变换不变态，被假定是唯一的——获得的有关态矢的希尔伯特空间组成的理论体系。现在让我们更为详细地来看一下位于这个体系之下的三个

假设。

第一个是定域性(locality)假设。根据狄拉克(1948)，“定域动力学变量是描述一个时空点上的物理条件的量。例如，场量和场量的导数”。并且，“如果能建立一个在其中所有动力学变量都是可定域的波函数的表象，那么量子理论中的动力学系统将被定义为可定域的”。在量子电动力学中，定域性假设作为粒子点模型的一份遗产和它对粒子之间相互作用的描述，采取了在一个类空平面上相互对易(玻色子)或反对易(费米子)的场算符的形式，这保证了在相关的类空位置上场量的测量能够相互独立地进行。 [211]

第二个是场算符假设。当约当在1925年、狄拉克在1927年把量子力学的方法扩展到电磁学时，电磁场分量就被从经典的对易变量提升为量子力学算符。相同的程序也能用于那些描述费米子的场(Jordan and Klein，1927；Jordan and Wigner，1928)。按照与粒子相联系的量子的产生和湮灭算符，这些定域场算符有一个直接的物理解释，它们被看作是场的局域激发或退激。根据不确定性原理，对于粒子的产生而言，严格的局域激发隐含着可以获得任意数量的能量和动量。因此把场算符应用到真空态的结果，不是获得一个包含一个单粒子的态，而是获得一个包含任意数目粒子的叠加态，这个叠加态只受相应的量子数守恒的限制。[1]从物理上讲，这意味着与微粒论不同，在量子电动力学中相互作用机制下的本体，实质上是场而不是粒子。从数学上讲，这意味着算符场由其全体矩阵元定义。现在可以明白，这些占压倒性比例的矩阵元指涉远离实验经验的能量和动量。

第三个是裸真空的充盈假设。有许多论据反对充盈真空的假设，其中最强有力的一个论据建立在协变推理之上：由于真空一定是零能量、零动量、零角动量、零电荷，无论什么都为零的洛伦兹变换

不变态，因此它必定是一个空无一物的态（Weisskopf，1983）。然而，当分析某些据猜想是由真空涨落造成的现象时，同样还是那些反对充盈真空假设的物理学家，却又心照不宣地把真空看作某种具有物质的东西，即看作一种可极化介质，或假设它是一种潜在的基质或激烈活动的舞台。换言之，他们实际上接纳了充盈真空的假设。

总之，考虑到在量子电动力学中，相互作用的粒子和传递相互作用的介质分别是费米子场量子和电磁场量子，即它们都是场本体的显示，因此，量子电动力学应该被看作一种场论。然而，量子电动力学中的定域场不是通常意义上的场。作为定域量子化的场，它们在很大程度上已失去了自身的连续性。量子场的这一新面貌被证明对量子场纲领中的相互作用机制的表述有深刻影响。

[212] 由于量子电动力学中的相互作用是由场量子和被称作离散虚量子（而不是连续场）的传递之间的局域耦合实现的，因此这个表述通过局域耦合的概念，深深地植根于算符场局域激发的概念之中。从而，在量子电动力学中，相互作用的计算必须考虑包含任意高能量的虚过程。然而，除了诸如幺正性那样的通常限制所强加的影响以外，就这些能量而言，本质上不存在相信这一理论是正确的经验证据。从数学上看，把这些在任意高能量上的虚过程包括进来，导致了明显不可确定的无穷大量。因此，发散困难是内在的，而不是外在的，是量子电动力学的真正性质。它们在量子场论的正则表述中是必要的。在这个意义上，发散的产生清楚地指出了在未重正化量子电动力学的概念结构中深刻的不连贯。或用另一种方式说，任何量子场论只有当它可重正化时，才能被认为是一个连贯的理论。正如我们将要了解的，这就是为何量子场纲领的进一步发展会受到限制的原因。

8.2 失败的尝试

20 世纪 40 年代后期，在量子电动力学的重正化微扰表述取得显著成功以后，量子电动力学被当作核弱相互作用理论和核强相互作用理论模拟的一个范例。但是，把相同的方法应用到核相互作用的尝试是失败的。这些失败解释了为什么在 20 世纪 50 年代物理学家对于量子场论的兴趣衰减了。

费米最初在 1933 年提出旨在描述 β 衰变的弱相互作用理论，这一理论在 20 世纪 50 年代被加以修正，以包括宇称破坏，并以有质量中间矢量介子(W 介子)理论的形式被重新表述[2]，因此它享有与量子电动力学相同的理论结构。虽然量子电动力学被看作是可重正化的，但是龟渊(Kamefuchi)在 1951 年的论文中指出，费米的四费米子直接相互作用是不可重正化的。稍后又有研究证明，即使 W 介子理论也是不可重正化的。这里给出这一缺陷的原因。

对于一个可重正化的理论来说，当我们定位在高阶时，原始发散图类型的数目必须保持有限，以使无穷大量能被吸收进有限多个参量中，诸如质量和耦合常数，它们是任意参量，能由实验确定。然而，在费米的原始表述(四线耦合)中，空间耦合常数 $G_W(\sim m^2)$ 隐含着，G_W 的更高次幂与更严重的无穷大量和更多类型的发散积分相联系。因此，要求有无限多的任意参量来吸收无穷大量。在 W 介子理论中，耦合常数是无量纲的，但是，有质量的矢量介子传播子 $q^\mu q^\nu/m^2q^2$ 的大动量渐近形式却包含量纲系数 m^{-2}。由于在微扰理论的更高阶上的 m^{-2} 因子必须由越来越多的发散积分补偿，因此有质量的矢量介子传播子的大动量渐近形式是不可重正化的。 [213]

对于强核力的介子理论的情形，特别是在 π 介子和核子的赝标量耦合形式中，情形更为复杂。形式上讲，赝标量耦合是可重正化

的。然而，由于这个可重正化形式的耦合常数太大，以至于不允许使用微扰理论——重正化程序起作用的唯一框架——因此它的可重正化性是不可实现的。更为特殊的是，重正化过程实际上只在如下这样的微扰方法中才有可能，即在涉及微扰计算的每一步上从一个“无穷大”项巧妙地减去另外一个“无穷大”项。微扰方法本身决定性地依赖于耦合常数相对很小这一事实。于是，情况立即变得显然，微扰方法不适用于介子理论，介子理论相应的耦合常数比 1 要大得多（$g^2/hc \approx 15$）。量子电动力学类型的量子场论的这一失败尝试，为色散关系方法的流行铺平了道路，也为丘(Chew)的 S 矩阵理论方法的采纳铺平了道路。在这一失败尝试中，量子场论的整个框架受到相当多的理论学家的反对(见 8.5 节)。

但是，即使在电磁学领域中，量子场论的失败也是明显的。费尔德曼(David Feldman)在他 1949 年的论文中谈到，矢量介子的电磁相互作用是不可重正化的。在 1951 年，彼得曼(Peterman)和斯蒂克尔伯格发现，磁矩与电磁场的相互作用(形如$f\bar{\psi}\sigma_{\mu\nu}\psi F_{\mu\nu}$的泡利项)是不可重正化的。稍后，海特勒(1961)和其他人发现，除了电荷以外完全等同的粒子(诸如 π 介子和核子)的质量差，是不能用重正化理论计算的。不难证实，如果质量差具有电磁性起源，那么发散的电磁自能将导致无穷大的质量差。这个困难清楚地表明，重正化理论不能实现泡利关于重正化理论将提供说明“基本粒子”质量比的普适理论的希望。

除了这些失败外，量子场论的重正化表述也被批评为是一个过于狭窄的框架，以至于不能容纳对诸如 CP 破坏的弱相互作用和引力相互作用这种重要现象的描述。但是，重正化理论最严重的缺陷从 20 [214] 世纪 60 年代末以来才变得明显，其时人们认识到，重正化理论与飒

刺性地作为重正化要求的结果而出现在量子场论高阶上的手征性反常和迹反常，有直接和不可调和的冲突(见 8.7 节)。

量子场论的这些失败在物理学家中间产生了一种危机意识，它要求物理学家澄清，对于量子场论的可重正化性，什么才是适当的态度。

8.3 对于可重正化性的各种态度

20 世纪 40 年代以来，物理学家面对的一个基本问题是，是否自然界中所有相互作用都是可重正化的，进一步的问题是，是否只有可重正化的理论才是可接受的。

戴森意识到这个问题的答案是否定的，并以这种观点在奥尔德斯顿会议上作了报告(Schweber，1986)。在戴森关于量子电动力学的可重正化性的经典论文发表以后，这种立场受到立即出现的具体的、清楚的、否定的例子的支持(见 8.2 节)。

对于其他物理学家来说，例如贝特，由于他们已把可重正化性从量子电动力学的一个属性提升到了指导理论选择的一个调节性原理，因而这个问题的答案是肯定的(见 Schweber，Bethe and Hoffmann，1955)。他们依据预测力(predictive power)为他们的立场辩护。他们的论据是：既然基本物理学的目标是用公式表述拥有相当大预测力的理论，那么"基本定律"一定只包含有限数量的参量。只有可重正化的理论才与这一要求一致。虽然不可重正化理论的发散项能通过被吸收进适当指定的参量中而得以消除，但是这要求有无穷多参量，而且这样的理论最初要用一个出现在拉格朗日函数中的无穷多参量来定义。根据他们的可重正化性原理，与电磁场相互作用的自旋为 1/2 的带电粒子的相互作用拉格朗日函数，不可能包含泡利力矩。运用相同的推理，费米的弱相互作用理论失去了它作为一个基本理论

的地位。重正化另一个应用上的限制是对强相互作用中耦合 π 介子与核子的赝矢量的拒斥。

然而，重正化理论自身的内在连贯性受到了戴森、卡伦(Källen)、朗道(Landau)和其他人的挑战。卡伦在1953年的论文中声称，从所有重正化常量是有限的假设出发，能推得量子电动力学中至少有一个[215] 重正化常量一定是无穷大的。若干年来，大多数物理学家把这个矛盾的结果理解为量子电动力学不连贯的证据。然而，正如后来一些批评家所指出的那样(例如，Gasiorowicz *et al*., 1959)，卡伦的结果依赖于一些众所周知的不可靠的论据，这些论据涉及在无穷多态上求积与求和顺序的相互交换，因而是非决定性的。卡伦(1966)本人后来承认了这一含混性。

挑战重正化理论连贯性的一个更为严重的论据，是用微扰理论的崩溃表达的。众所周知，戴森的重正化理论只在微扰理论的框架内得到系统表述。微扰重正化理论的产物是一组对场论的格林函数形式上定义得很好的幂级数。然而，人们不久意识到这些级数(特别是适合 S 矩阵的级数)极可能发散。因此，理论学家陷入混乱的状态，不能给出这个问题的答案：场论的微扰级数在什么意义上定义了一个解？十分有趣的是，第一个对微扰重正化理论幻灭的物理学家是戴森本人。1952年，戴森给出了一个有独创性的论证，表明重正化之后所有的幂级数展开都是发散的。随后由赫斯特(Hurst, 1952)、蒂林(Thirring, 1953)、彼得曼(1953a, b)、贾菲(Jaffe, 1965)以及其他公理化场论(axiomatic field theory)和构造场论(constructive field theory)的理论学家给出的讨论，进一步加重了如下断言：绝大多数重正化场论的微扰级数是发散的，尽管在大多数情形中尚不存在完全的证明。

对于格林函数而言，一个发散微扰级数可能仍然是这个理论的一个渐近解答。20 世纪 70 年代中期，构造场论学家建立了某些场论模型的解，这些人后验地指出，这个解由其微扰展开唯一地确定(Wightman, 1976)。然而，这些解只适合于二维或三维时空连续统中的场论模型。就更现实的四维量子电动力学而言，赫斯特在 1952 年已提出，量子电动力学与实验符合得极好，这说明微扰级数可能是一个渐近展开式。

然而，具有讽刺意味的是，卡伦、朗道，特别是盖尔曼和洛(Gell-Mann and Low, 1954)，对于量子电动力学的高能行为的研究表明，在量子电动力学中，作为电荷重正化的一个必然结果，微扰方法不可避免地失效了。朗道及其合作者进一步主张，微扰框架中的余留或者将导致毫无相互作用(零重正化电荷)[3]，或者将导致使得理论明显不连贯的鬼态(ghost states)的出现(Landau, 1955; Landau and Pomeranchuck, 1955)。两种结果都表明，微扰理论在重正化量子电动力学中不适用。

在种类繁多的非阿贝尔规范理论中，特别是在量子色动力学(QCD)中发现了渐近自由之后，人们表达了对于微扰量子色动力学将去除朗道鬼态、进而消除对于量子场论连贯性的大部分怀疑的希望。然而，这种希望没有持续多久。人们不久就意识到，在高能处消失的鬼态会在低能处重新出现(Collins, 1984)。因此，事实强有力地不断提醒着场论学家微扰理论的应用限制。结果，大体上说，量子场论的连贯性问题，特别是它的微扰表述的连贯性问题，仍然处于不确定的状态。 [216]

对可重正化性问题，理论物理学家的态度截然不同。对大多数职业物理学家来说，连贯性只是一个学究式的问题。作为实用主义

者(pragmatists)，他们只受他们的科学经验的指引，很少有兴趣去思考一个理论的终极连贯性。

朗道和丘采纳的立场更为激进和极端。他们拒斥的不仅仅是相互作用的特殊形式和量子场论的微扰描述，而且拒斥量子场论本身的一般框架(见8.5节)。对他们而言，定域场算符概念本身，和对在微观时空区域中相互作用的任何具体机制的假定，完全是不可接受的，因为这些都太过思辨，以至于甚至在原则上也是不可观测的。他们的立场得到了量子场论存在发散且重正化理论缺乏连贯性这样一些证据的支持，尽管朗道对于重正化量子电动力学的不连贯性的论据还不能称为是决定性的。

施温格关于重正化的观点特别有趣，这不仅仅因为他是重正化理论的奠基人之一，而且主要因为他对重正化纲领的基础给出了有洞察力的分析，并且是重正化纲领最为尖锐的批评者之一。根据施温格，采用定域场算符作为其概念基础的未重正化描述，包含了对高能细部敏感的物理粒子的动力学结构的猜测性假设。然而，我们没有理由相信在高能领域这一理论是正确的。克拉默斯认为，量子场论应该有一个不依赖于结构的特征，施温格赞同这一规则并把它接受为一个指导原则，他详细阐述的重正化程序去除了任何所涉及的极高能量过程、相关的小距离以及内部结构的假设。因此，他促进人们把注意力从关于局域激发和相互作用的假说性世界，转移到关于物理粒子的可观测世界。

[217] 然而，施温格发现，先引入与物理无关的结构假设，仅仅在最后删除它们以获得物理上有意义的结果，以这种迂回方式前行，是不可接受的。这个批评构成了对重正化的哲学拒斥。但是，如果定域算符场论能产生有意义的结果，那么重正化在该理论中将是重要的和不可

避免的。为了使批评产生逻辑结果，施温格(1970，1973a，b)引入了用数字估值的(非算符)的源和数字场来取代定域场算符。这些源象征着构成对于物理系统的测量的干涉。而且，相关场的所有矩阵元、算符场方程和对易关系式，都能用这些源来表达。此外，人们业已表明，作用量原理能给出整个形式体系的简洁表达。

根据施温格，他的源理论把有限量看作基本的，因而免于发散。这个理论也是充分可延展的，能归并新的实验结果，并能以一种合理的方式外推它们。最重要的是，这一理论能这样做而没有跌入必须把理论延伸到任意高能量的陷阱，这构成了未知的新物理学势必面临的未开发的领域。

因此，从施温格的观点出发，重正化的最终命运将是被消除，从对自然的任何描述中被排除出去。他试图通过放弃定域算符场的概念，进而彻底改变量子场论的基础来实施这一点。施温格研究进路(其基础是在他 1951 年的论文中打下的，在 20 世纪 60 和 70 年代他又做了详细阐述)的激进特征，直到 20 世纪 70 年代中期才被认可，其时可重正化性原理开始受到挑战。到那时为止，运用重正化群方法进行的研究探明了重正化和可重正化性的新洞见，导致了对于重正化和量子场论的新理解，也普遍导致了对于科学理论的全新看法。因此，人们重新燃起了对于不可重正化理论的兴趣，"有效场论"(effective field theory)方法开始获得它的声望(见 11.4 节)。在这种变化了的概念语境中，一些敏锐的理论学家开始意识到，施温格的思想对于基本物理学前景的根本性改变是重要的(Weinberg，1979)。

8.4 公理化方法

虽然量子场论在 20 世纪 50 年代开始逐渐衰退，但是它并没有死

亡。有一些物理学家对于量子场论采取了积极的态度。公理化场论[218]学家的态度最为积极[4]，他们竭力阐明量子场论的数学基础，希望清除其最显而易见的不连贯性。

在希尔伯特传统的精神鼓舞下，公理化场论学家试图用公理化来解决量子场论的内在连贯性问题，把这当作给出概念问题的清楚解答的唯一途径。

当希尔伯特尽力用由数学实体组成的形式体系的连贯性证明，使数学实体的使用合法化时，那些公理化场论学家从相反方向进行包抄。他们通过建构非平凡的例子——这些例子的存在只是公理的一个推论——来尽力证明量子场论的内在连贯性。他们没有对量子场论基础作激进的改变，而只是尽力逐步克服关于量子场论的连贯性的明显困难。虽然依然有许多重要问题遗留了下来，但是他们再也找不到哪里还有量子场论包含基本不连贯性的丝毫迹象。

对于公理化场论学家而言，定域场并非完全定域。他们把定域场看作赋值算符，并且只在分布的意义上存在。也就是说，定域场只能被理解为赋值算符分布，它们能用在无穷大处快速衰减的无穷大可微测试函数来定义，或者也能用具有坚实支持力的测试函数来定义。因此，对于公理化场论学家而言，询问在时空点 P 上的场值是没有意义的。在这种方法中重要的是在 P 的一个如此小邻域内场的模糊值，我们希望让这个值作用在其支持力包含在被选择的邻域内的测试函数上。从本质上讲，这是对于修正精确的点模型的物理思想的一个数学表示。考虑到如下事实：量子场论中的发散起源于点模型，而重正化的实质是吸收无穷大量，这等价于把精确的点模型模糊化，因此，在公理化场论中模糊化(smearing)的概念似乎表明，在公理化场论的基本动机和重正化理论之间有一种深刻的联系。

对于公理化场论学家来说，由于紫外发散的存在和源于无穷大时空体积的无穷大量，一个未重正化理论毫无疑问是不连贯的。这种不连贯必须从理论中清理出去，并且要把福克表象(Fock representation)作为对正则对易关系的外尔形式(Weyl form)的一个候选者排除出去。这两种无穷大量的出现使得定义一个哈密顿算符成为不可能，量子场论的正则量子化的整个方案坍塌了。

重正化理论中的连贯性问题完全不同于未重正化理论中的连贯性问题。紫外发散被设想为可以通过重正化程序规避。公理化场论学家声称，一些遗留的困难，诸如如何定义定域场和它们的等价类，如何规定渐近条件和连带约化公式，可以借助于分布理论和赋范代数，以一种严密的形式进行分析(Wightman，1989)。 [219]

当然，以这种方式定义的理论在微扰理论的意义上仍然是不可重正化的。然而，从20世纪70年代中期以来，人们已付出了巨大的努力，用构造场论的方法来理解不可重正化理论的结构，并建立使一个不可重正化理论可以有意义的条件。这项研究的一个显著结果是：一些不可重正化理论的解，只有有限数量的任意参量。这与它们按照微扰级数的描述相反。人们已猜想，在微扰理论中要重正化无穷多参量的必要性，可能来自一个不合理的幂级数展开式(Wightman，1986)。

确实，在这些努力中，公理化场论和构造场论的理论学家已展示了一个开放的和可变通的心境。然而，理解重正化理论的基础和证明其连贯性的进一步发展，可能涉及一些还没有被现有理论的任何公理证明为正确的并被吸纳的假设上的改变。无论如何，尽管有将近40年的精深努力，然而构建一个可解的四维场论的失败，却表明公理化场论学家和构造场论学家在希尔伯特的意义上解决量子场论的连

贯性问题，有相当大的困难。这也多少挫败了他们最初的乐观主义。

一个有趣但还没有被挖掘的论题是，在 1955 年到 1956 年期间，公理化场论在构造 S 矩阵理论的发展中所起的作用。根据莱曼、西曼齐克和齐默尔曼（Lehmann，Symanzik and Zimmermann，LSZ，1955，1957）以及怀特曼（Wightman，1956），量子场论或者它的 S 矩阵元，能直接只用格林函数（或推迟函数，或场乘积的真空期望值）表达。正如怀特曼（1989）曾经指出的，格林函数的 LSZ 约化公式“是不可数色散理论计算的出发点”；按照推迟函数，对于 $2 \to n$ 个粒子反作用力来说，LSZ 约化公式“对于 $n=2$，是证明色散关系的出发点”。值得注意的是，南部阳一郎在对色散关系作出重要贡献（1957）之前，在 1955—1956 年期间曾发表了两篇关于量子场论中格林函数的结构的论文（1955，1956）。这些论文富有公理化场论的精神并为双重色散关系的发展铺平了道路，这对于 S 矩阵理论的概念发展是至关重要的一步（见 8.5 节）。

[220] 8.5 S 矩阵理论

正如上文提到的，20 世纪 50 年代，量子场论中的危机意识很重。除了在用公式表示弱相互作用和强相互作用的可重正化理论上的失效以外，在 20 世纪 50 年代中期，朗道及其合作者（1954a，b，c，d；1955）指出，除了非相互作用粒子以外，定域场论似乎没有解。他们指出，只有把无穷大项从幂级数中除去，可重正化微扰理论才会起作用，并且对四维时空中任何量子场论来说，通常没有非平凡的、非微扰的幺正解。虽然可以产生正规的幂级数解并满足展开式的每一阶上的幺正性，但是级数本身可能不是收敛的。他们还声称已证明

量子电动力学必须有零重正化电荷。[5]朗道把这些概念困难当作证据，用来证明量子场论作为一个建立在量子定域场算符、微观时空、微观因果性和拉格朗日方法等概念基础上的研究框架是不充分的。他竭力主张量子场论的不连贯似乎在这些不可观测量中有其根源。朗道所提出的用来反驳量子场论的这一论据，尽管不被其时的大多数物理学家认真地接受，但是对于一些激进的物理学家来说的确有强烈的哲学吸引力。

量子场论中的另一个概念困难与其原子论思维有关。量子场论的一个基本假设是，由场方程描述的基本场与“基本粒子”相关联，基本粒子反过来被认为是整个宇宙的基本建筑砖块。在 20 世纪 30 年代早期，尽管有 β 衰变的存在，但人们已经意识到，中子不应该被认为是一个质子和一个电子的复合体。因此，物质分割为基本的建筑砖块，只能被认为是一个近似的模型，在非相对论性情形中很成功，但是在高能物理学中不再适用。不可能断定质子和中子哪一个是更为基本的粒子，它们中的每一个不得不被看作同等基本。同样的思想不得不被应用到在 20 世纪 50 年代后期和 60 年代早期发现的大量新粒子。人们直觉上很清楚，这些新粒子不可能全部都是基本的。因此，量子场论被迫要在基本粒子和复合粒子之间作出一个清楚的划分。直到 20 世纪 60 年代，关于这一主题的研究文献才开始出现，但是并没有给出令人满意的划分标准。[6]因此，在强相互作用的范围内，用最先偶然发现的、被看作基本的少数粒子来探究量子场论，似乎是不得要领的。

此外，在 20 世纪 60 年代初期，量子场论也被打上了经验上不充分的烙印，因为它没有给出与雷杰轨迹(Regge trajectory)相关联的粒子家族存在的任何解释。与此相关的其他困难是：量子场论似乎不 [221]

能产生解析散射辐，并且似乎也在色散关系中引入了任意的减去法参量。

与这一情境相反，在色散关系的发展中出现了一种新趋势，丘和他的同事把雷杰极点(Regge pole)和靴袢假说协调地结合在一起，形成一个高度抽象的方案——解析S矩阵理论(SMT)；在20世纪50年代和60年代，这一理论成为强相互作用领域中一个主要的研究纲领。

S矩阵理论最初由海森伯在1943年和1944年提出(1943a，b；1944)，作为一种取代量子场论的自洽的研究纲领。海森伯对于量子场论的批评集中在发散困难上；发散困难提醒他，未来的理论还需夯实基础。他关于S矩阵的工作的目的是想在量子场论的基础上提取那些普遍适用的、因此将包含在未来的正确理论中的概念。洛伦兹变换不变性、幺正性和解析性[7]是这一物质世界的概念，并被具体表现在S矩阵理论中。S矩阵由渐近态确定，因此是由实验直接给出的量。而且，S矩阵理论似乎不遭遇任何发散困难。事实上，有限性(finiteness)是海森伯研究纲领的主要目标，而重正化纲领的出现排除了这一目标的必要性。与之前遇到的一些其他困难一起，重正化使海森伯的S矩阵理论退居到了幕后。然而，由这一纲领提出的问题，诸如"因果性和S矩阵元的解析性之间的关联是什么？"(Jost，1947；Toll，1952，1956)和"如何从散射数据中确定薛定谔理论的相互作用势？"，导致盖尔曼和戈德伯格(Gell-Mann and Goldberger，1954；with W. Thirring，1954)提出色散理论纲领(dispersion theory programme)。[8]

色散理论纲领为计算S矩阵元提供了动力学方案，这在海森伯纲领中是缺乏的。戴森计算S矩阵元的方法(1949a，b)，虽然也受海森

伯纲领的启示，但在本质上是与海森伯纲领非常不同的一种方法，因为它明确地涉及场论。盖尔曼、戈德伯格和蒂林提出色散纲领的初始动机是从场论中提取精确的结果。他们所遵循的散射幅的解析性和色散关系，源于对可观测场算符的微观因果性的假设。他们用到的另一个“原理”，即交叉对称性(crossing symmetry)，被看作一个普遍的性质，被表征场论的微扰展开式的费恩曼图所满足。

在1956年召开的罗切斯特会议上，盖尔曼(1956)声称，如果合适的边界条件被加在动量空间的无穷大动量上，那么交叉性、解析性和幺正性将确定散射幅。他进一步声称，这对于确定一个场论也几乎是充分的。因此，色散理论纲领被其发明者之一盖尔曼认为只是一种在质壳(包括虚动量在内)上明确表达场论的方式。情况确是这样，虽然在罗切斯特会议上盖尔曼也提到，如果对色散纲领进行非微扰处理，那么将令人回想起海森伯的希望——直接写下S矩阵而不是从场论中计算它。 [222]

事实上，在色散纲领中获得的结果是相当独立于场论细节的。幺正性将任一散射幅的虚部与涉及散射幅的平方的一个总截面关联起来，因果性(通过解析性)把散射幅的实部和虚部相互关联起来。以这种方式，一个闭合的、自洽的非线性方程组产生了(这或者能以微扰的方式解决，或者能以某种非微扰的方式解决)，这使物理学家能从基本原理中获得完整的S矩阵，并且能从量子场论的概念困难和计算困难中把它们拯救出来。正是色散纲领和海森伯纲领之间的这样一种内在逻辑关联，使得盖尔曼对于海森伯S矩阵理论的“偶然”提及，成为20世纪50、60年代解析S矩阵理论进一步发展的强有力的推动力。[9]

然而，作为一个动力学方案，色散理论缺乏一个至关重要的因

素：一个关于“力”的观念的对应物是什么样的清楚概念。在这个理论中，计算方案是在质壳上实行，因此只涉及渐近态，其中，粒子外在于彼此的相互作用区。因此，能使这个方案描述相互作用的一个关键性问题是，如何定义一个表征散射幅的解析函数。由盖尔曼、戈德伯格和蒂林(1954)明确表达的初始色散关系，清楚地展示了粒子极点，这对于定义一个解析函数在数学上是必要的，因为否则，如果用刘维尔定理来定义，这个解析函数将是一个令人不感兴趣的常数。然而，对于理论的这一特征在最初并没有给出任何强调，更不用说对S矩阵的奇异性所扮演的角色的准确理解了，即对一个不同于量子场论的新纲领在提供“力”的概念上的作用没有适当的理解。

在这个方向上的进展是通过以下三种方式取得的：(i)把戈德伯格的相对论性色散关系(1955a，b)与丘和洛的π介子—核子散射理论(1956)结合在一起；(ii)把盖尔曼和戈德伯格的交叉对称性(1954)和曼德尔施塔姆(Mandelstam)的双重色散关系(1958)结合在一起；(iii)把解析延拓从线性扩展到角动量(Regge，1958a，b；1959；1960)。让我们依次考察这些进展。

[223]

(i) 丘和洛的理论把丘用于π介子—核子散射的静态模型(1953a，b)与洛用于散射幅的非线性积分方程(1954)联合在一起，并且只涉及重正化量。在单介子近似中，它有一个满足幺正性和交叉性的解，有一个单极点和好的渐近行为，但它遭受了CDD歧义性(CDD ambiguity)的折磨。[10]力的观念在传统的拉格朗日场论中是明确的。然而，丘和洛的表述通过应用解析函数，允许与实验数据有更直接的联系。解析函数的极点是这一联系的关键。极点的位置与粒子质量联结在一起，极点处的留数与“力的强度”(耦合常数)联结在一起。这导致关于力的一种新概念：力存在于一个解析S矩阵的奇点中，并

使得这个半相对论性模型变为完全相对论性的(Chew, Goldberger, Low, and Nambu, 1957a, b)。 而且，色散关系本身被理解为用奇异性表达一个解析S矩阵元的柯西—黎曼公式。[11]

(ii) 交叉性是盖尔曼和戈德伯格在他们的一篇论文(1954)中发表的；交叉性最初被理解为一种关系，通过这一关系入射粒子变为出射反粒子。 曼德尔施塔姆的双重色散关系(1958)是从能量变量到散射角的解析延拓，或者同曼德尔施塔姆变量一起是从 s(在 s 信道中的能量变量)到 t(在 t 信道中的不变能量，或者通过交叉，在 s 信道中的动量传递变量)的解析延拓。 曼德尔施塔姆双重色散关系的一个重要结果是，它把交叉性的概念转变为一种新的动力学方案，在这个动力学的方案中，在一个给定信道中的力存在于交叉信道的一个极点之上。

这个新方案反过来导致了用于理解强子的靴袢方法。 双重色散关系使得丘和曼德尔施塔姆除了能分析 π 介子—核子散射外，还能分析 π 介子—π 介子散射(1960)。 他们发现，两个 π 介子(后来被命名为 ρ 介子)的一个自旋为1的束缚态构成了一种力，这种力经由交叉性是制造相同的束缚态的中介；也就是说，作为一种“力”的 ρ 介子产生作为一个粒子的 ρ 介子。 靴袢概念作为这种计算的一个结果而引入。 没过多久，丘和弗劳奇(Chew and Frautschi, 1961a, b)就提出，如在 ρ 介子情形中一样，所有的强子都是其他强子的束缚态，这些强子在交叉信道中由强子的极点表征的“力”支撑。 所有强子通过这样一种靴袢机制而自生，可望用来自洽地和唯一地确定它们的全部性质。 [224]

曼德尔施塔姆表象被证明与势散射理论中的薛定谔方程等价。[12] 而且，研究表明，薛定谔方程能作为对低能处的色散关系的一个近似

而获得。[13]这就支持了如下猜测：一个具有确立得很好的非相对论性势散射理论的自洽纲领，在它的极限情况，可以取代量子场论。在这个纲领中，动力学不由时空中一个具体的相互作用模型规定，而是由受制于最大解析性要求的散射幅的奇异性结构决定。散射辐的奇异性结构要求在这个散射幅中没有不同于由幺正性和交叉性要求的奇异性出现。[14]在这些靴袢计算中所遗留的一个更深的困难由渐近条件问题提出。对于势散射，这一困难被雷杰(Tullio Regge)解决。

(iii)在证明位势理论中的双重色散关系时，雷杰(1958a，b；1959；1960)能证明，同时把S矩阵延拓到复能量(s)平面和复角动量(j)平面是可能的。因此，在j平面中一个特殊极点的位置是s的一个解析函数，而一个固定的极点($\alpha(s)$ = 常量)是不被允许的。如果在某一能量($s>0$)上，Re $\alpha(s)$的值经过一个正整数或零，那么在这一点上将有一个自旋等于这个整数的物理共振态或束缚态。因此，通常，在j平面中一个单极点的轨迹作为改变了的s，对应于一个不同的js和不同质量的粒子家族。然而，当$s<0$时[15]，雷杰证明，雷杰轨迹$\alpha(s)$控制弹性散射幅的渐近行为，它与$t^{\alpha(s)}$成正比，当s充分为负时，$\alpha(s)$也为负。因此，只是在那些$l<\alpha(s)$的分波中，才可能有束缚态。这隐含着能在t信道的渐近行为中探测到$s<0$的轨迹，而最为重要的是，这隐含着折磨量子场论的发散困难至少是可以避免的。

这些结果被丘和弗劳奇(1960，1961c)注意到，并被假设在相对论性情形中也是正确的。最大解析原理从线性动量("第一种"动量)到角动量("第二种"动量)的合成延拓(resulting extension)隐含着，信道中的力存在于交叉信道的移动奇点上，这就使得对靴袢假说作如下重新表述成为必要：所有强子，作为S矩阵的极点，位于雷杰

轨迹上。

靴袢假说的这一变种为动力学强子模型提供了基础。在这个也称作“核民主”(nuclear democracy)的模型中，所有的强子都被看作或者是复合粒子或者是组分粒子，或者是结合力，这取决于它们涉及的特定过程。人们根本关心的不是粒子，而是它们的相互作用过程，强子的结构问题被按照强子反应幅的结构重新表述。[16] [225]

根据丘和弗劳奇，区分复合粒子和基本粒子的判据是一个粒子是否依赖于雷杰轨迹：复合粒子与一个在 j 平面上作为 s 的一个函数移动的雷杰极点相联系，而基本粒子极点仅与角动量相联系，不承认在 j 平面上有延拓。根据这个判据，他们声称：

> 如果自然界中确实出现了 $j = 1/2$ 和 $j = 0$ 的基本粒子极点，那么可以认为，直接用 S 矩阵工作只是一个评估传统场论的技巧。另一方面，如果对所有重子和介子极点承认在 j 平面上的延拓，那么关于强相互作用的传统场论不仅是不必要的，而且是明显误导的或许甚至是错误的。
>
> (Chew and Frautschi, 1961a)

这是对量子场论的一个严重挑战，盖尔曼和其他场论学家对这一挑战的回应，导致了对调和量子场论和 S 矩阵理论的一次尝试，即雷杰化纲领(Reggeization programme)。[17]

作为一个反对量子场论的基本框架，S 矩阵理论是由极少数粒子理论学家在 20 世纪 50、60 年代贡献的一个结果。然而，其中心思想和原理主要由朗道和丘提供。朗道为 S 矩阵理论奠定基础的贡献包括：(i)主张量子场论不足以处理相互作用；(ii)促进人们认可在强

子和解析S矩阵的极点之间存在普遍的对应性[18]；(iii)表述S矩阵的奇异性的图解法和规则。[19]朗道的图解法不同于费恩曼图，它绝不等价于微扰理论：所有相关粒子对一个物理过程的贡献都被包括在奇异性中[20]，图中的线条与表征渐近态的物理强子相对应。不必考虑重正化。因此，表征物理过程的朗道图成为一个新的研究主题，也成为S矩阵理论的基础的主要部分。[21]事实上，朗道关于图解法的论文(1959)成了20世纪60年代发展的公理化解析S矩阵理论的出发点。[22]

然而，丘是引领这场运动远离量子场论、朝向S矩阵理论发展的集成者和真正领袖。除了在建立新的动力学方案的整个三阶段中的[226]原创性贡献以外，丘也提供了一些哲学论据来支持强子的这一动力学模型。他充满热情地给出反驳量子场论中采用的原子论范式的论据，并拒斥任意指定基本粒子的思想。受盖尔曼和朗道思想的影响，也受他自己关于雷杰极点和靴袢计算的研究的激励，丘在写《强相互作用的S矩阵理论》(*S-Matrix Theory of Strong Interactions*, 1961)一书期间，于1961年曾在拉霍亚召开的一次讨论会上作了一次演讲；在这次演讲中，他简要说明了他的思想和海森伯的旧S矩阵理论思想之间的联系。在这一著名的演讲中，丘切断了与量子场论的联系，而采纳了S矩阵理论。[23]丘采取了极端激进的立场，坚持认为在强相互作用物理学中，我们不得不放弃量子场论而使用S矩阵理论。在1962年1月在纽约召开的美国物理学学会会议上，丘(1962a)宣称他致力于解析S矩阵理论，而拒斥被他“最熟悉和最亲密的朋友”所持有的“场论是一种同样合适的表达方式”的观点。在他看来，“基本的强相互作用概念，在纯S矩阵方法中是简单和优美的，而对场论来说，如果不是不可能的，就是怪诞的”。

丘的反场论立场主要基于两个假设：(i)核民主和靴袢假说；(ii)最大解析性。根据(i)，所有的强子都是复合的，所有物理上有意义的量都是由幺正性方程和色散关系自洽地和唯一地确定，不允许有任意指定的基本量(诸如与出现在量子场论中的基本粒子或基本场相关的量)。最大解析性的假设除了要求由朗道规则要求的最小集合，或者要求所有的强子均位于雷杰轨迹上以外，还要求S矩阵没有奇异性。其时，所有强子位于雷杰轨迹上这种性质，通常被理解为复合粒子的特征。根据丘，雷杰假说提供了实现靴袢思想的唯一途径。他坚持量子场论与最大解析性原理不相容，因为量子场论假定了存在不能产生雷杰极点的基本粒子。由于低质量的强子，特别是核子，被发现位于雷杰轨迹上，因此，丘就能与把质子看作基本粒子的传统信念分道扬镳，并声称所有的强相互作用物理学将从最大解析性、幺正性和其他S矩阵原理中产生。

对丘而言，量子场论之所以是不可接受的，主要因为量子场论假定了基本粒子的存在，这不可避免地导致与核民主和靴袢假说相冲突，也与解析性相冲突。基本粒子的不存在，对于丘的反场论立场是那么关键，以至于他甚至选择这一判据来刻画其新理论的特征。[24] 除了丘支持靴袢假说和最大解析性的论据以外，丘拒斥基本粒子的概念也为西岛和彦(Kazuhiko Nishijima，1957，1958)、齐默尔曼(1958)以及哈格(Rudolf Haag，1958)的早期工作所支持。这些作者声称，就散射理论而言，复合粒子和基本粒子之间不存在任何差别，因为所有的粒子产生相同的极点和截止，与它们的起源不相关。丘甚至能诉诸费恩曼原理——正确的理论不应该为确定哪些粒子是基本的提供可能性。[25]因此，从1961年6月的拉霍亚会议到1962年7月的日内瓦会议，丘给场论学家施加了巨大的压力。

[227]

作为靴袢假说的一个原创者和主要鼓吹者，丘把自洽性和唯一性原则作为S矩阵理论的哲学基础。在20世纪60年代前半期，丘强有力地影响了强子物理学共同体。他的哲学立场，尽管甚至被他的一些亲密合作者(如盖尔曼和洛)视为是教条的和迷信的[26]，然而却被他及其合作者在S矩阵理论框架内获得的一些得到很好证实的物理结果所支持。这些结果中最令人信服的是，人们发现大量重子和介子共振子几乎位于线性雷杰轨迹上。靴袢动力学另一个给人深刻印象的证明是ρ共振子的质量和宽度的计算，人们发现计算结果接近于实验值。[27]丘关于核子、π介子和(3，3)共振子的倒易靴袢计算，为这种靴袢计算提供了一个更为复杂的例子(1962a)。

S矩阵理论在物理学上及其哲学诉求上的成功，使得它在20世纪60年代早期非常流行。在强子动力学方面工作的物理学家，不可避免地要对丘向量子场论的挑战作出回应。然而，除了丘的一些最忠实的追随者之外，甚至在他的合作者之中，也很难发现有谁是全心全意地忠于丘的激进立场的。不过，仍然有一些物理学家，例如布兰肯贝克勒(Blankenbecler)，试图采纳对丘的立场来说是至关重要的思想，他们实际上成为了丘的立场的支持者。其他人，如盖尔曼，虽然吸收了一些丘鼓吹的思想，但是拒绝丘关于S矩阵理论与量子场论相对立的观点，试图调和这两种方法，甚至不惜付出概念上不连贯的代价。[28]

无论如何，几乎每一位粒子物理学家都强烈地感受到了丘施加在粒子物理学基础概念上的压力。例如，在20世纪60年代前半期，进行了关于靴袢内在对称性的雄心勃勃的研究。[29]另一个启发性的例子与温伯格有关，他后来成为一名杰出的场论学家。在1964—1965年间，温伯格(1964a，c；1965b)试图把S矩阵理论从强子相互作用扩

展到电磁相互作用[30]和引力相互作用，并在推导规范不变性和麦克斯韦方程组上取得了成功，也在S矩阵理论框架内成功推导出等效原理和爱因斯坦方程。温伯格的目的是“质问场论在理解电磁学和引力上的必要性”，因为直到那时，电磁学和引力一直被认为是场论最可靠的例子。其时温伯格“还不清楚” [228]

> 场论是否将继续在粒子物理学中起作用，或者说是否它最终将被纯粹的S矩阵理论所取代。
>
> (Weinberg，1964a)

温伯格在分析中使用费恩曼图的表达方式时，他请求读者“在此，请欣赏我们孩童时代训练的结果，而不是对场论有任何本质上的依赖”(同上)。

温伯格的态度反映了当时流行的气氛，这一时代的气氛一年后由戴森刻画如下：

> 现在许多人极度怀疑场论与强相互作用物理学的相关性。场论处于防御现在流行的S矩阵的处境中……容易想象在若干年内场论的概念将完全从高能物理学的日常工作词汇中剔除出去。
>
> (Dyson，1965)

量子场论和S矩阵理论在基本假设上的不同可总结如下：

(i) 在S矩阵理论中，基本本体是过程而不是实体；量子场论的建筑砖块是基本场或它们的量子——基本粒子。

(ii) 在S矩阵理论中，力由奇点结构表示；而在量子场论中，力由虚量子的传播表示。

(iii) 在量子场论中，费米子数守恒，玻色子数不守恒；而在S矩阵理论中，费米子和玻色子之间的这一差异不那么明显。

(iv) 在S矩阵理论中只有复合粒子，它们的参量可由动力学方程唯一地确定；而在量子场论中存在具有任意参量的基本粒子。

(v) S矩阵理论产生没有任何固定奇点的解析振幅；量子场论产生有固定奇点的振幅，这起源于有固定自旋的基本粒子。这一区分导致了它们高能行为的差异，S矩阵理论中的高能行为要比量子场论中的柔和得多。

(vi) 从方法论上讲，S矩阵理论单独从可观测量开始对可观测事件的计算，没有引入任何不可观测的实体。系统的动力学由普适原理确定，没有引入任何不可观测的机制。结果，不需要重正化。然而，量子场论从基本场(有时是不可观测的，如在夸克和胶子的情形中)和用来确定基本系统的动力学的规范群开始计算。量子场论给出微观时空中相互作用的详细描述。这里，指导原理是经验上的"试错"(trial and error)。由于没有获得前后一致的解，因此，为了获得可与实验相比较的结果，不得不求助于微扰方法和重正化程序。

[229]

(vii) 在S矩阵理论中用靴袢机制实现的统一(unification)是横向的；在量子场论中这种统一是纵向的，即向下还原和向上重建。

强子物理学中S矩阵理论的流行，在20世纪60年代中期开始衰落，部分是因为在处理截止和试图把这一理论扩展到两体信道之外

时，遇到了数学上的困难，部分是因为那时夸克的观念正在流行。然而，从概念上讲，按照它与量子场论相比较的基本假设，S 矩阵理论仍然在较长一段时间里很有效。靴袢思想，它要求在 t 信道和 u 信道中互换的强子集等价于在 s 信道中形成的共振子集和束缚强子集，这导致双共振模型(dual-resonance model)。反过来，这又隐含着强子的弦绘景并为超对称和超弦理论的引入提供了背景。

从历史上看，自从 S 矩阵理论产生以来，量子场论作为一个基本框架，一直处于来自 S 矩阵理论的概念压力之下，正如后者也处于来自前者的压力之下一样。一些物理学家试图把量子场论还原为 S 矩阵理论的一种极限情形[31]，另一些物理学家则把 S 矩阵理论的原理看作量子场论的属性。[32]温伯格把量子场论看作实施 S 矩阵理论的公理的一个方便途径(1985；1986a，b)。然而，量子场论和 S 矩阵理论作为两种独立的研究纲领，几乎没有物理学家去研究它们之间的相互影响；这种研究对于我们理解粒子物理学的概念发展，可能证明是有帮助的。[33]

8.6　部分守恒轴矢流假说[34]和流代数

对于量子场论危机的第二种类型的回应，是建立在对称性考虑的基础之上的。对称性方法首先被应用于强子的弱相互作用和电磁相互作用领域。后来，它被扩展到低能强相互作用领域。从现象学的角度看，这些相互作用似乎可根据强子流(hadron currents)用有效哈密顿函数很好地描述。人们推测，强子流是从强子场中构造出来的，然而，在某种程度上，它的细节只能有待于涉及强子的性质、支配强子行为的动力学定律(场方程)和解释这些定律的理论这些基本事件来决定。在 20 世纪 50、60 年代，物理学家对除了电磁学之外的 [230]

精确定律知之甚少，不能解答任何已提出的解释强子动力学的现实模型所提出的问题。因此，在主流物理学家中间，通常放弃沿着量子电动力学线路来发展强相互作用理论的努力，尽管杨振宁和米尔斯给出了一个提议，并且随后出现了一些对这个提议的回应。在那个时期，在强子物理学中与S矩阵理论一起出现的第二种趋向是，先不急于分析来自"第一原理"的强子流。相反，这种趋向把注意力集中在对称性上，而对称性被认为可以独立于动力学细节来推断其隐含的内容。

在20世纪50年代后期、60年代早期，一个被称为部分守恒轴矢流假说(PCAC hypothesis)和流代数的新研究纲领逐渐成形。在这个纲领中，定域场论的对称性被用来超越这个理论的动力学细节，并拥有超出最初几阶微扰的有效性。这个对称性纲领(symmetry programme)，与解析性纲领S矩阵理论一样，也是一个准自洽纲领。它之所以是一个准自洽的纲领，首先是因为其起源于量子场论。它是从量子场论中提取出来的，并且能用场论模型来研究。这个纲领研究的主要对象——流算符(current operators)，被设想原则上是从定域场算符中构造出来，并能以与在量子场论中场算符相同方式进行运算；流算符能被嵌入两个物理态，以表示它对初态的作用，这一作用造成初态向末态的跃迁。由于对称性纲领是从一个没有动力学源(场方程)的物理系统的一些普遍属性(这个纲领中的对称性与S矩阵理论中的解析性、幺正性等相对照)中提取物理信息，实现了一个独立的理论推进，因此是准自洽的。这个纲领的出发点是对称性。流主要被认作是对称性的表象，而不是从动力学系统中推导出来的，其时动力学系统是不可解的。

如果我们关于物理世界的知识，能从一个展现在各种典型的拉格

朗日函数和物理流中的先验数学对称群中推导出来，那么柏拉图主义者和数学神秘主义的信奉者，可能已从这一纲领中找到了某种支持。但这是一个误解，或者说，至少是一个幻想。相反，这个纲领的真正性质是以逻辑经验为特征的。不变群结构，诸如 U(1)，SU(2)，SU(3)等等，绝不是先验的，相反，它们是由实验数据中显现出的精确的或近似的规律性来间接表明的。流的矩阵元被猜想为描述了真实粒子中发生的物理上可观测的过程。

[231] 这样一种现象学方法，拥有通常与逻辑经验主义方法论相联系的所有优点和弱点。积极的一方面，坚守直接可观测量(S 矩阵元、形状因子等)，有助于它规避在一个可作重正化处理的定域场中由于引入不可观测的理论结构所造成的全部困难；代数关系的形式操作(流对易式等)，不但简化了真实情形，使得它们可操作，而且呈现出普遍有效性。消极的一方面，对于通常的不可观测实体和过程的否定，特别是对微观动力学的否定，妨碍它朝着对基本粒子及其行为的深入理解的方向前进。而且，如果没有对于动力学的适当理解，就不能保证经常倾向于将真实情况过分简化的纯形式操作是普遍有效的。

正如我们将在这节的稍后部分详细考察的，我们会有趣地注意到，对称性纲领的演化很快就在 1967 年左右到达了极限。不但其理论上的潜能几乎已被耗尽，而且人们发现它一些基于形式操作的预测与实验直接相冲突。实验令人信服地表明，对称性纲领的基本假设之一，即流对易式是不依赖于动力学细节的，是完全错误的。为了阐明这一情形，人们付出了巨大的努力，不久就导致了关于流代数的场论研究。其结果是对于定域场论中的反常(anomalies)的深刻发现(见 8.7 节)。这一偶然发现正好击中了重正化观念的核心。不但规

范反常的出现破坏了对称性，因而破坏了理论的可重正化性，而且更重要的是，反常的概念为正确理解重正化群方程提供了基础（见 8.8 节），也为重正化的新概念提供了重要的组成部分（见11.4 节）。对称性纲领这一具有讽刺意味的事是引人注目的。对称性纲领的出发点，是在排除动力学考虑的意义上对定域场论的拒斥。然而，为了批判性地考察这个纲领本身，其发展却为回到这一场论框架铺平了道路。当然，这不是简单地返回到以前的情形，而是作为一个否定之否定，回到了一个更高的层次，对于定域场论和重正化的丰富结构有了一个更为深刻的理解。

流及其矩阵元；形状因子

探究动力学系统的对称性，而又完全不求助于动力学，这样的思想在20 世纪50 年代是完全无效的，然而，从50 年代末起，基本物理学中物理流的普遍出现，首先让人想到了这一思想。在数学上，经由诺特（Noether）的工作，众所周知的事实是，守恒流遵循动力学系统的对称性，并构成了一个对称群表象。然而，对称纲领研究的流不是纯粹的数学对象。确切地说，它们首先起因于对弱相互作用和电磁相互作用的描述。只是后来在研究这些流的过程中，随着守恒流或近似守恒流假设的作出，人们才把物理流和对称流视为一体，于是，动力学系统的对称性得以充分利用。

[232]

电磁学为物理学家的推理提供了原型。由于电荷守恒，轻子的或者强子的电磁流保持守恒，因此能被看作与 U(1)对称流等同，而 U(1)对称流是与一个物理矢量玻色子场耦合的。在这种情形中，守恒流既是系统的对称性的表象，也是相互作用过程中能量交换的介质。

继电磁流的模式之后，经过一些必要的修改，物理学家关于弱流（weak current）的概念有所发展。格尔什坦（S. S. Gershtein）和泽利

多维奇(Y. B. Zel'dovich)在他们的论文(1955)中，首先提及并放弃了同位旋守恒矢量流(和相应的矢量玻色子)与β衰变具有某种关联的思想。施温格在他1957年的论文中从对称性考虑入手，更仔细地思考了矢量玻色子的同位旋三重态的存在，这些矢量玻色子的普适耦合将给出弱相互作用和电磁相互作用：有质量的带电Z粒子，将以与无质量光子作为介质传递电磁相互作用相同的方式，传递弱相互作用。这是弱相互作用的中间矢量玻色子理论的开始，具有深远的影响和重大的意义。

然而，在理解弱流的本质和结构上取得的更为实质性的进展，是受到弱相互作用中宇称不守恒的理论分析和实验证实的强有力进展的直接启发。在费米的原始论文中，β衰变是用矢量(V)流的形式书写的。1957年对宇称不守恒的观测要求一个轴矢量(A)耦合。苏达尔尚和马尔沙克(E. C. G. Sudarshan and Marshak, 1958)，还有费恩曼和盖尔曼(1958)，通过假设宇称的最大破坏，提出弱流的V-A形式。费恩曼和盖尔曼也提出了一个普适费米理论，在这一理论中，有效的弱相互作用哈密顿函数被写成弱流的乘积形式。依照格尔什坦和泽利多维奇所放弃的思想，还有施温格更为成熟的思想，费恩曼和盖尔曼猜测，弱流的矢量部分与电磁流的同位旋矢量部分同一；众所周知，电磁流的同位旋矢量是守恒的。因此，他们提出在物理弱流和SU(2)同位旋对称性之间有某种联系。布卢德曼(Sidney Bludman)在他1958年的论文中更进一步地详细阐述了这种联系，从同位旋SU(2)对称性扩展到手征SU(2)对称性[35]，并从这一对称性推出弱流。 [233]

布卢德曼的理论实质上是在弱相互作用的(杨—米尔斯型)规范理论方面的一个尝试；在这一规范理论中，有质量的带电矢量场与守恒

弱流耦合。樱井纯(J.J.Sakurai)在他1960年的论文中，提出了一个与弱相互作用理论有相似理论结构的有影响的强相互作用理论。在樱井的理论中，强相互作用由三个守恒矢量流描述：同位旋流、超荷流和重子流。这些流与重矢量玻色子耦合。然而，在自发对称性破缺思想融入量子场论之前[其时自发对称性破缺思想刚刚开始成形(见10.3节)]，规范玻色子的质量不可能得到说明，这个理论不久就退化为矢量—介子支配模型。(在第三篇有对非阿贝尔或杨—米尔斯型规范理论的更多讨论。)

在所有这些情形(施温格；费恩曼和盖尔曼；布卢德曼；以及樱井)中，包含在电磁、弱和强相互作用中的强子流，被认为是定义明确的洛伦兹四维矢量，是从基本强子场构造出来的，但不能用一种可靠的方式分析。这些流能够相互作用，它们的耦合能通过交换矢量量子实现。最为重要的是，它们的发散和正则对易式表达了系统的对称性。因此，通过以一种合乎规则的方式巧妙地处理这些流，系统的对称性似乎是可利用的。假设其为真，这是对称性纲领思想的基础。

在耦合常数中，最低阶的强子流(或电荷)的矩阵元是直接可观测量，它们服从色散关系，被认为受一个极点的支配。撇开容易确定的运动学因子，通过运用对称性原理，这一矩阵元的剩余部分能用四维动量变换的一些未知的标量函数，即所谓的形状因子，诸如电磁形状因子 $F_i(q^2)$ $(i=1, 2, 3)$ 来表达；这一电磁形状因子通过以下公式定义：

$$\langle p' \mid {}_a j_\mu^{em}(0) \mid p \rangle_b = u(p')_a [F_1(Q^2)(p'+p)_\mu + F_2(Q^2)(p'-p)_\mu + F_3(Q^2)\gamma_\mu] u(p)_b, \qquad (1)$$

其中 $q^2 = (p' - p)^2$；弱形状因子 $f_v(q^2)$，$f_m(q^2)$，$f_s(q^2)$，$g_A(q^2)$，$g_P(q^2)$和 $g_E(q^2)$由以下公式定义： [234]

$$(2\pi)^3 \langle p' |_{s'p} V_\mu(0) | p \rangle_{sn} = iu(p')_{s'p}[\gamma_\mu f_v(Q^2) + \sigma_{\mu\nu} Q_\nu f_m(Q^2) + iQ_\mu f_s(Q^2)] u(p)_{sn} \quad (2)$$

和

$$(2\pi)^3 \langle p' |_{s'p} A_\mu(0) | p \rangle_{sn} = iu(p')_{s'p}\{\gamma_5[\gamma_\mu g_A(Q^2) + iQ_\mu g_P(Q^2) + i(p'+p)_\mu g_E(Q^2)]\} u(p)_{sn}, \quad (3)$$

其中，V_μ 和 A_μ 分别是弱矢量流和弱轴矢流，而 $q_\mu = (p' - p)_\mu$。

这些弱(或电磁)形状因子被认为包含了有效强相互作用对基本弱(或电磁)相互作用的修正的所有物理信息。不存在预测这些形状因子的量化细节的任何方式，因为这些是由强相互作用的动力学确定的，在 20 世纪 50、60 年代，它们是不可获得的。然而，这些形状因子仍然能够得到有效的研究。第一，对称性考虑给它们施加了强有力的限制，限制它们的数量并在它们中间建立联系。第二，极点支配假说有助于就虚粒子对形状因子的贡献提供具体的建议。例如，为了解释电子—质子散射形状因子，南部阳一郎(1957a)首先推测了 Ω 介子的存在，然后在 1962 年，实验发现证实了这一推测。第三，由于形状因子是可观测量，因此所有基于对称性(和/或解析性)考虑的分析，均可由实验检验。因此，对称性纲领可被看作一个完全合理的研究纲领。现在，让我们从历史视角来对这一纲领作贴近的考察。

守恒矢量流(CVC)假说

费恩曼和盖尔曼在他们 1958 年的论文中推测，作为核 β 衰变原因的强子弱流(J_μ)的矢量部分 V_μ 是守恒的。他们的初始动机是要

对核β衰变中的矢量流耦合常数 $G_V(=Gf_V(0))$ 与 μ 介子β衰变中的矢量流耦合常数 G 很接近这一观测结果，给出一个解释。这一观测结果意味着，对矢量流耦合常数而言，重正化因子 $f_V(0)$[36] 非常接近于1，并且不被强相互作用重正化。[37] 一个对强相互作用守恒的流必定是对称流。也就是说，它可以从强相互作用的一个对称性中推导出来。因此，费恩曼和盖尔曼很自然地就把β衰变的流 V_μ 和同位旋矢量流看成是同一的。这明确地意味着，在β衰变中产生电荷的电性矢量流 V_μ 与同位旋的上旋流 $J^V_{\mu+}=J^V_{\mu1}+iJ^V_{\mu2}$ 同一，而 V^+_μ 将与同位旋的下旋流 $J^V_{\mu-}=J^V_{\mu1}-iJ^V_{\mu2}$ 同一，而电磁流 $j^{em}_\mu(=j^s_\mu+j^v_\mu)$ 的同位旋矢量部分(j^v_μ)是相同的同位旋矢量流的第三分量 $J^V_{\mu3}$。

[235]

显然，这个假说使得已知的电磁过程与某些弱过程通过同位旋旋转关联起来成为可能。例如，由于对一个处于静止状态的π介子

$$\langle\pi^+|J^V_{03}=j^{em}_0(I=1)|\pi^+\rangle=1, \tag{4}$$

对π介子β衰变来说，相应的S矩阵元必定为

$$\langle\pi^0|V_0|\pi^+\rangle=2^{1/2} \tag{5}$$

因此，弱矢量流的标度是固定的。另一个例子是核β衰变的矢量部分的形状因子，它与中子和质子的电子散射的矢量部分的形状因子之间的关系为[38]：

$$F_i(Q^2)^+=F_i(Q^2)^P-F_i(Q^2)^N \tag{6}$$

从这些关系中，盖尔曼(1958)给出了一个关于"弱磁性"$f_m(0)$的明确预测：

$$\begin{aligned}f_m(0)&=F_2(0)^P-F_2(0)^N=(\mu_P-\mu_N)/2M\\&=(1.79+1.91)/2M=3.70/2M,\end{aligned} \tag{7}$$

这随后为 Y. K. Lee，L. W. Mo 和吴健雄所证实(1963)。

守恒矢量流假说在粒子物理学史上占有突出位置。第一，虽然没有解释，但它指出了在弱、电磁和强相互作用之间存在着一些联系，进而揭示了它们的幺正性特征：物理上的弱流起源于强相互作用的 SU(2)同位旋对称性，与在相同的对称性操作下的电磁流相联系。第二，既然矢量流守恒表示一个守恒定律，因此守恒矢量流假说隐含着杨—米尔斯的思想可以从强相互作用领域扩展到弱相互作用领域。这一扩展随后在布卢德曼关于弱相互作用的规范理论中，用一个手征 SU(2) × SU(2)对称性予以具体表达。事实上，守恒矢量流假说的成功增强了物理学家对于杨—米尔斯理论的信念；作为一个正反馈，也 [236] 鼓励樱井提出了一个扩展的强相互作用的规范理论。第三，弱矢量流 $J^V_{\mu i}$ 的守恒隐含着相伴弱电荷 Q_i^V 是 SU(2)弱同位旋对称性的生成元。它们产生一个满足如下正则对易关系式的 SU(2)李代数：

$$[Q_i^V, Q_j^V] = i\varepsilon^{ijk}Q_j^V, \quad i, j, k = 1, 2, 3 \tag{8}$$

方程(8)是流代数中一个重大方案的起点。它处理存留奇异性的电性弱矢量流的电荷和电磁流的同位旋矢量部分的电荷。这对于整个电性弱矢量流(包括奇异性变化项)和整个电磁流的扩展是简单明了的，是以强相互作用(Gell-Mann and Ne'eman，1964)的近似 SU(3)对称性[39]，或更成熟的卡比玻模型(Cabibbo，1963)为基础的。在这种情形中，所有的弱矢量流和电磁矢量流，包括同位旋标量电磁流，同位旋矢量电磁流的三重态和弱奇异性守恒流，以及弱奇异性变化流，都属于相同的八重态，相伴电荷 Q_α^V 产生一个更大的 SU(3)李代数，满足：

$$[Q_\alpha^V, Q_\beta^V] = i f_{\alpha\beta\gamma} Q_\gamma^V, \quad \alpha, \beta, \gamma = 1, 2, \cdots, 8, \tag{9}$$

其中 $f_{\alpha\beta\gamma}$ 是 SU(3)结构常量。关于弱轴矢流 $J^A_{\mu\alpha}$ 和相伴电荷 Q^A_α，也给出了一个相似的提议。在手征对称性限制中，一个手征 $SU(3)_L \times SU(3)_R$ 代数是用附加的关系：

$$[Q^A_\alpha(x_0), Q^A_\beta(x_0)] = \mathrm{i} f_{\alpha\beta\gamma} Q^V_\gamma(x_0), \quad (10)^{40}$$

和

$$[Q^V_\alpha(x_0), Q^A_\beta(x_0)] = \mathrm{i} f_{\alpha\beta\gamma} Q^A_\gamma(x_0) \quad (11)$$

假设的。

当盖尔曼在论文(1962a)中第一次系统地提出流代数的框架时，他建议这些代数关系式应该用于补充弱流和电磁流的矩阵元计算中的色散关系。然而，在考察流代数及其应用之前，我们必须先来查看另一个假说——部分守恒轴矢流假说，尽管该假说在概念上多少是独立的，但它实际上经常与流代数缠绕在一起。

[237] **部分守恒轴矢流假说**

如果弱流 $J_\mu(= V_\mu - A_\mu)$ 的矢量部分 V_μ 因强相互作用而守恒，那么对于轴向部分 A_μ 来说情形又将如何呢？对于轴向耦合常数来说，重正化因子 $g_A(0)$[41] 的实验值接近于 $1(g_A/f_V \approx 1.25)$。费恩曼和盖尔曼(1958)猜测，如果这个值精确地等于 1(他们认为实验上不排除会有这样的结果)，那么对于轴矢流耦合常数将没有重正化，轴矢流也将守恒。因此，他们提议尝试构造守恒轴矢流，并考察所涉及的对称群的含义。

他们的提议不久就被波尔金霍尼(John Polkinghorne，1958)和泰勒(John C. Taylor，1958)采纳，也为布卢德曼(1958)和居尔塞伊(Feza Gürsey，1960a，b)更为系统化地吸收。布卢德曼和居尔塞伊的工作有一个强对称性倾向，并以施温格(1957)和图切克(Bruno

Touschek，1957)的工作为先驱。因此，对于轴矢流的研究，如同对于矢量流的研究一样，受对称性考虑的支配，这一研究后来导致了流代数的形成。

然而，沿着这一路线的进展受到了严格限制，因为轴矢流的精确守恒受到了几个实验事实的阻拦。第一，从单位1出发的 g_A/f_V 表明，轴向耦合存在重正化效应。第二，正如戈德伯格和特赖曼(Goldberger and Sam Treiman，1958a)所指出的，轴矢流守恒将引入较大的有效赝标量耦合，这与实验相矛盾。第三，正如波尔金霍尼所指出的，它隐含着存在K介子宇称的双重线。第四，正如泰勒所指出的，它隐含着一个等于零的 π 介子衰变常量，这也与实验相矛盾。

研究轴矢流的真实动力实际上来源于另一个理论传统，即解析性传统。当对守恒轴矢流的推测似乎是徒劳时，戈德伯格和特赖曼发表了他们关于 π 介子衰变的色散关系的论文。在他们的计算中，π 介子衰变常量 f_π 的吸收部分被看作是色散变量 q^2（离壳 π 介子质量的平方)的一个函数，并涉及从离壳 π 介子到一个中间态(这主要是核子—反核子对的态，通过轴矢流与轻子耦合)完全集的过渡。最终结果为：

$$f_\pi = Mg_A / g_{\pi NN} \qquad (12)$$

其中 M 是核子的质量，g_A 是在核 β 衰变中轴矢耦合常数，$g_{\pi NN}$ 是 π 介子—核子强相互作用耦合常数。除了与实验符合得极好之外，戈德伯格—特赖曼关系式(12)还展示了强相互作用量和弱相互作用量之间的关系。这一显著的成功与对称性或流守恒无关，它从有说服力的假说开始，通过对此假说提供一个解释或推导，挑战了理论物理学 [238]

家。这也鼓舞了沿着不同方向进行研究，最终导致了部分守恒轴矢流假说的产生。

在寻求更为基本的假说的努力中，南部的方法(1960d)占据着突出的位置，因为它建立在对于对称性破缺的一个更为深刻的理解的基础之上。南部主张，如果一个有关无质量核子场 ψ 的理论，在 $\psi \to \exp(i\alpha \cdot \tau\gamma_5)\psi$ 变换下，具有不变性，那么轴矢流将守恒。然而，如果真空态在这一变换下保持不变，那么核子将变成有质量的，而对称性也将明显破缺。但是，在一个赝标量态中，由于一个无质量的束缚核子—反核子对[42]的出现(这可被看作 π 介子态与核子的耦合)，对称性将得以恢复。这种自发对称性破缺确保了 π 介子的存在，但不能解释 π 介子衰变：由于 π 介子的存在恢复了对称性，因而轴矢流仍然是守恒的。

然而，南部遵从了盖尔曼和利维(Gell-Mann and Maurice Levy，1960)的建议，并假定一个小的裸核质量为 π 介子质量量级。用这种方式，γ_5 对称性显然破缺，π 介子变得有质量，而且轴矢流的矩阵元能用形状因子 $g_A(q^2)$ 和 $G_A(q^2)$ 表示为：

$$\langle p' \mid A_\mu \mid p\rangle_{sn} = u(p')[ig_A(Q^2)\gamma_\mu\gamma_5 - G_A(Q^2)Q_\mu\gamma_5]u(p),$$
$$Q = p' - p \tag{13}$$

这里，$G_A(Q^2)$ 由 $q^2 = -m_\pi$ 处的极点支配，起源于核子和轴矢流之间交换一个虚 π 介子。也就是说：

$$G_A(Q^2) = 2f_\pi g_{\pi NN}/(Q^2 + m_\pi^2) \tag{14}$$

南部进一步主张，如果轴矢流几乎守恒(在当 $q^2 \gg m_\pi^2$ 时发散消失的意义上)，并且形状因子变化缓慢，那么戈德伯格—特赖曼关系式(12)可从方程(13)和(14)得出。

另一条流行的研究进路由盖尔曼及其合作者开创(Gell-Mann and Levy, 1960; Bernstein, Fubini, Gell-Mann, and Thirring, 1960)。他们研究场论模型，在诸如梯度耦合模型、σ模型和非线性σ模型上进行研究，并沿着泰勒的例子指引的方向构造了其发散与π介子场成正比的轴矢流：

$$\partial A'_{\mu}/\partial x_{\mu} = \mu^2 f_{\pi} \phi \qquad (15)$$

[239]

在这一方法中，他们所作的重要假说是：轴矢流的发散是一个柔和算子(gentle operator)。这意味着其矩阵元满足未减少的色散关系(π介子质量可变，就像其色散度可变一样)，并且变化缓慢；在低频处，这些矩阵元受到来自π介子中间态的贡献的支配，在高频处(当π介子的质量可忽略时)，这些矩阵元消失。因此可以说，轴矢流几乎守恒。

就部分守恒轴矢流假说而言，这两种方法差不多等价。然而，它们的理论基础很不相同。南部的方法是建立在对自发对称性破缺(SSB)理解的基础之上。无质量π介子的出现是自发对称性破缺的一个结果，尽管它也有助于恢复整个系统的对称性；而在盖尔曼的方法中没有这样的理解，尽管消失的π介子质量被用来表明某种近似对称性。

现在应该清楚，柔和假说(gentleness hypothesis)——部分守恒轴矢流假说的实质——在概念上不依赖于任何对称性考虑。它的重要性在于它允许轴矢流的矩阵元物理上从π介子质壳延拓到离壳这一事实。当离壳π介子质量变量趋于零时，经常会产生令人感兴趣的论断。这一假说给我们以希望：当π介子质量变量返回到物理质壳上时，这些论断可能或多或少保持为真；而且这提供了一个重要的预见

力，并能为实验所证实。

实际上，由于牵涉到低能 π 介子物理学中的复杂情况，部分守恒轴矢流假说的应用经常与流代数的思想缠绕在一起(见下文)。然而，有一些独立的检验。例如，阿德勒(1964)表明，中微子诱导的弱反应 $\nu + \mathrm{P} \to \mathrm{L} + \beta$(这里P代表质子，L代表轻子，$\beta$ 代表一个强子系统)的矩阵元 M，借助于矢量流守恒假说，能写成：

$$M_{q^2\to 0} \sim G \ \langle \beta \mid \partial A_\mu^{1+i2}/\partial x_\mu \mid \mathrm{P}\rangle \ , \ q = p_\nu - p_{\mathrm{L}} \qquad (16)$$

强 π 介子反应 $\pi^+ + \mathrm{P} \to \beta$ 的矩阵元 M_π，通过使用约化公式能写成：

$$M_\pi = lim_{q^2\to -\mu^2}[(q^2+\mu^2)/2^{1/2}] \ \langle \beta \mid \phi^{1+i2} \mid \mathrm{P}\rangle \qquad (17)$$

(其中 q 是 π 介子的动量，$\phi^{1+i2}/2^{1/2}$ 是 π^+ 场算符)，这两个矩阵元能以下面的方式相互关联(通过使用方程(15)并通过延拓 M_π 离开 q^2 处的质壳到 $q^2=0$ 处，这被猜想导致了一个与物理 π 介子振幅相似的振幅)：[240]

$$\mid M\mid^2_{q^2\to 0} \sim G^2 f_\pi^2 \mid M_\pi \mid^2 \qquad (18)$$

上述例子，像戈德伯格—特赖曼关系式一样，展示了弱相互作用和强相互作用之间的关系。事实上，部分守恒轴矢流假说导致了联结弱相互作用和强相互作用的一整套关系。如果人们知道强相互作用跃迁幅 $\mathrm{T}(\pi^+ + \alpha \to \beta)$，那么这些关系就允许人们去预测弱相互作用矩阵元 $\langle \beta \mid \partial_\mu J_\mu^A \mid \alpha \rangle$。通过使用这些关系，阿德勒证明(1965)，在只有玻恩近似贡献和 $\langle \beta \mid \partial_\mu J_\mu^A \mid \alpha \rangle$ 能用弱耦合常数和强耦合常数表达的确定情形中，才能消除弱耦合常数，并获得一个只涉及强相互作用的连贯性条件。阿德勒获得的 π 介子—核子散射的著名的连贯性条件是，在对称同位旋 π 介子—核子散射幅 $A^{\pi N(+)}$(核子 $K^{NN\pi}$ 的 π

介子的形状因子)和重正化 π 介子—核子耦合常数 g_r 之间存在着非平凡关系：

$$g_r^2/M = A^{\pi N(+)}(\nu = 0, \nu_\beta = 0, k^2 = 0)/K^{NN\pi}(k^2 = 0), \tag{19}$$

其中 M 是核子的质量，$-k^2$ 是初始 π 介子的质量平方，$\nu = -(p_1 + p_2)k/2M$，$\nu_\beta = qk/2M$，p_1、p_2 和 q 分别为初始核子、终了核子和终了 π 介子的四维动量。阿德勒证明在 10%的误差范围之内，这一关系与实验相一致。[43]

在部分守恒轴矢流假说各种成功的应用和预测之中，有一个突出的例外，即一组涉及中性 π 介子的衰变，其中之一是 $\pi^0 \to 2\gamma$ 衰变。正如萨瑟兰(D. G. Sutherland，1967)和韦尔特曼(1967)所指出的，部分守恒轴矢流假说隐含着衰变率为零，这与实验相矛盾。部分守恒轴矢流假说这一引起争论的失败，引发了对于部分守恒轴矢流以及流代数的理论基础的考察。结果导致了对于反常的清楚认识(见 8.7 节)。

流代数

1961 年盖尔曼首先提出，对于在强子的电磁和弱相互作用中产生的流的时间分量而言，流代数的本质要素是一组等时对易(ETC)关系。然而，对于在等时对易关系下形成一个封闭代数系统的电荷密度来说，相关的物理流必定是与物理系统的一些连续对称性相联系的诺特流(Noether currents)。20 世纪 60 年代初期，人们注意到了强 [241] 子物理学数据中的规律性。其中一些规律性能按照雷杰轨迹来说明，另一些则用对称性表达。在第二种情况中，我们能找到三种类型。第一种类型涉及精确的几何对称性，它隐含着能量、动量等的守恒，也隐含着 CPT 不变性(在时间反演、空间反射和电荷共轭联合

变换下的不变性）。在第二种类型中，我们发现精确的内部对称性，它与电荷、重子数等等的守恒有关。第三种类型是最令人感兴趣的情形，因为首先，它在那时就涉及了强相互作用的所谓的近似对称性[44]（诸如与同位旋和奇异性的近似守恒相联系的那些近似对称性），其次，它提供了导致流代数观念的思维的出发点。

在方法论上，流代数由其独立于动力学细节的特色所刻画。它有时会用到拉格朗日函数，但仅仅是为了表达由实验数据显示出的对称性，并产生服从于由这些对称性支配的正则对易关系的流。在数学上有两种方式来表达这些对称性。它们能用守恒定律来表示。然而，这种方法的效用仍然是很有限的。麻烦在于强子物理学中涉及的大多数对称性是近似对称性，而用这种方法不能解决确定它们有多近似的问题。这个问题也能并且经常被翻译成一个物理上更精确的问题，即如何固定这些流的矩阵元的标度问题，这个问题与近似对称性相联系，因而受强相互作用重正化的支配。在概念上，这个问题要求阐明物理学中近似对称性的意义和含义。

在从物理学上和概念上解决这个问题的一次尝试中，盖尔曼采纳了用李代数表示对称性的另一种方法；这种方法由对称性生成元组成，并在正则对易关系下是封闭的。

在历史上，盖尔曼的方法标志着在现代物理学中对易式使用的第三个阶段。在第一个阶段，1926 年海森伯引入位置 q 和动量 $p \equiv \delta L/\delta q$，$i[p, q] = (h/2\pi i)I$ 之间非相对论性的正则对易式，来表示量子化条件。这个对易式独立于拉格朗日函数 L 的特定形式，能用来推导独立于相互作用的有用结果，诸如托马斯—库恩加法定则（Thomas-Kuhn sum rule）。第二个阶段与蒂林用两个场算符的对易

[242] 式来表示微观因果性的方法有关，这两个场算符的对易式对于其函数

的自变量的类空分离应该等于零。 形式上，盖尔曼用对易式代数定义对称性的方式，相似于蒂林用对易式定义因果性的方式，因为两者都采纳了量子场论的当代表述。 然而，在精神上，盖尔曼更接近于海森伯而不是蒂林，因为对易式代数集注定会展示相对论性动力学中独立于相互作用的关系，这些关系能被用来提取物理信息，而无需求解这个理论的方程。

数学上，如果有一个线性无关算符 R_i 的有限集，并且任意两个 R_i 的一个对易式是 R_i 的一个线性组合：

$$[R_i(t), R_j(t)] = \mathrm{i} c_{ijk} R_k(t), \tag{20}$$

那么这个体系就叫做一个李代数。 物理上，这个抽象的公式在同位旋电荷算符中有它的现实化形式(参阅方程(8)和随后的讨论)。 心中有了守恒矢量流假说和同位旋对称性的想法，盖尔曼在 1962 年的论文中提出了一个 SU(3) 代数，又在 1964 年的论文中提出了一个 SU(3) × SU(3) 手征代数，它由矢量电荷和轴矢流八重态以及能量密度中的对称破缺项 u_0 组成(见方程(9)，(10)和(11))[45]，由此开创了一个重大的流代数研究纲领。

从概念上讲，在对盖尔曼纲领的理解上存在困难。 首先，在那时还没有被设想为与弱流耦合的矢量玻色子，因此流的作用显得非常神秘。 其次，流被假定为 SU(3) × SU(3) 对称性的表象。 然而，除了 U(1) 规范对称性之外，其大多数亚对称性都严重破缺。 因此，关于这些近似对称性的真实意义是什么，存在理解上的混乱。

盖尔曼对于这些困难的解决，既是根本的又是精巧的。 首先，他以一种抽象的方式考虑了更高的近似对称性，并把流看作研究的主要题目。 因此，矢量玻色子尽管对物理过程的形象化是有帮助的，

但不是必要的。其次，流是按照它们的等时对易关系而不是对称性本身来刻画的。这是在刻画流的这两种方式之间的一个细微然而却是关键的差异。形式上，当等时对易关系被看作正则场变量的函数时，只依赖于流的结构。如果拉格朗日函数中的对称破缺项不涉及微商耦合，因而不依赖于细节，那么流将保留最初的结构，等时对易关系也将保持不变。用这种方式，矢量和轴矢流，甚至是在强相互作用的SU(3)对称性破缺的情形下，不是在它们的电荷守恒的意义上，而是在它们的电荷满足SU(3)的等时对易关系的意义上，仍然有SU(3) × SU(3)对称性的八重态变换属性。因此，按照现在能被看作强相互作用的一个确切属性的等时对易关系，给近似对称性概念指定了一个虽是抽象然而却是精确的意义。

[243]

盖尔曼关于近似对称性的新颖理解，允许人们从一个破缺的对称性推导出可测量的量之间的精确关系。这些关系有助于确定流的矩阵元的标度，特别是有助于把轴向流的标度和矢量流的标度联合在一起。此外，就轻子和强子弱相互作用的强度而言，普适性假说也获得了准确的意义。也就是说，关于总的轻子弱流的等时对易关系，与关于总的强子弱流的等时对易关系是相同的。借助于包括 u_0 在内的代数关系，也能计算恒等式的导数，即重正化效应。因此，在盖尔曼的经典论文(1962a，1964a，b)发表以后，物理学共同体获得了一个可行的研究纲领。

在这个纲领中，中心主题是涉及 π 介子的过程，因为在手征限制($m_\pi \to 0$)下，π 介子的矩阵元能通过直接应用等时对易关系，在轴矢流发散的情况下结合 π 介子极点支配的部分守恒轴矢流概念来进行计算。例如，通过部分守恒轴矢流假说和归约公式，过程($i \to f + \pi$)的矩阵元与下列表达式发生关系

$$\langle f \mid \partial^{\mu} A_{\mu}^{a} \mid i \rangle = \mathrm{i}k^{\mu} \langle f \mid A_{\mu}^{a} \mid i \rangle \tag{21}$$

同样，包含两个软 π 介子的过程也能通过分析以下矩阵元来研究

$$\langle f \mid \mathrm{T}\partial^{\mu} j_{\mu}^{a}(x) \partial^{\nu} j_{\nu}^{b}(0) \mid i \rangle \tag{22}$$

在脱离微分算符的过程中，我们除了获得一个双重流矩阵元的双重发散外，还获得了一个双重流的等时对易式，后来证明这对于计算是重要的。

然而，直到 1965 年富比尼(S. Fubini)和富兰(G. Furlan)提出了追随这一纲领的某些技巧，关于这一纲领的研究才有了快速的进展。他们的中心思想是(i)在对易式的矩阵元中插入一个中间态的完全集，(ii)分离出一个核子的贡献，(iii)在无穷大动量系中计算。通过运用本质上相同的技巧，阿德勒和韦斯伯格(William Weisberger)分别独立地获得了一个用质心能量 W 处的零质量 $\pi^{+}(\pi^{-})$—质子散射总横截面来表达轴向形状因子 g_A 的求和定则：

[244]

$$1 - 1/g_A^2 = [4M_{\mathrm{N}}^2/g_{\mathrm{r}}^2 K^{\mathrm{NN}\pi}(0)^2(1/\pi)\int_{M_{\mathrm{N}}+M_{\pi}}^{\infty} W[\sigma_0^{+}(W) - \sigma_0^{-}(W)]\mathrm{d}W/(W^2 - M_{\mathrm{N}}^2) \tag{23}$$

这个公式产生了一个预测值 $|g_A| = 1.21$，该值与实验数据 $g_A^{\mathrm{exp}} \approx 1.23$ 非常一致。

阿德勒—韦斯伯格求和定则也能解释为关于 π 介子—核子散射的低能定理，其影响是巨大的。它提供了对于部分守恒轴矢流—流代数纲领的最著名的检验，因此，为这一纲领确立了一个范例。正如特赖曼曾经表述的那样，它的发表为这一纲领开创了黄金时代。在两年之内，大约有 500 篇论文紧随其后发表，它在低能定理和/或高能求和定则中是成功的。

然而，在无数的成功应用之中也有突出的失败。最著名的事例是：(i)关于 $\pi^0 \to 2\gamma$ 衰变率等于零的萨瑟兰—韦尔特曼定理；(ii)卡伦—格罗斯(Callan-Gross)求和定则(1969)，它预言在深度非弹性极限内，脱离质子的总纵向电致截面等于零。对这些失败的全面考察，与对某种形式操作的有效性的精心关注一起，在流代数和部分守恒轴矢流假说的语境中导致了对反常对称性破缺的深刻理解。这一理解令人信服地表明，不诉诸潜藏在代数关系下面的动力学的形式操作是不可靠的。它也表明，等时对易关系不依赖于相互作用的细节的假设是完全错误的，因而削弱了整个研究纲领的基础。

然而，在转向这一重要的发展之前，对这个纲领进行一些评论是适宜的。第一个评论关系到在这个纲领初次出现时就被察觉到的观点。因为在这个纲领中的流被猜想为既是物理上的弱流和电磁流，也是强相互作用的对称流，所以用弱和电磁过程似乎可能检测强相互作用的对称性。就味对称性(味对称性是质量简并度的一种表达)而言，这一期望已部分实现。

在这点上，就流代数和杨—米尔斯理论之间的可能关系而言，初步的成功引发了一些更高的期望。例如，盖尔曼(1962a；1964a，b；1987)声称，流代数证明味手征 SU(3) × SU(3) 对称性作为一种强相互作用的对称性是合理的，它将为强相互作用的杨—米尔斯理论产生一个规范群，至少等时对易关系是杨—米尔斯理论的一个先决条件。[245] 然而，它证明味对称性与一个杨—米尔斯类型的强相互作用理论的构造不相关，虽然它的确与强相互作用的低能动力学有某种相关性。人们也期望(Bernstein，1968)等时对易类型的代数约束条件将有助于挑选出流的动力学模式。这一期望同样也没有实现。

然而，在轻子物理学中，流代数的确在推进杨—米尔斯类型的理

论中起到了启发性的作用。通过用强相互作用的对称性把电磁流和弱流联系起来，它促进了在轻子物理学中引入新的量子数，即弱同位旋和弱超荷。由于与这些量子数相关联的弱流和电磁流被猜想为与矢量玻色子耦合，因此这些概念构造的确有助于构造一个杨—米尔斯类型的电弱理论。

第二个评论关系到方法论。部分守恒轴矢流—流代数纲领有意忽略了动力学细节，因此使得复杂的情形变得简单。正是这一简化使得在理解强相互作用的某些低能方面成为可能，这种理解主要是在 π 介子物理学上，这与强相互作用中的味手征对称性的自发破缺相关联。然而，当研究更为深入时，这种方法的不足很快就显露出来，1967 年后，人们逐渐认识到了说明这一纲领的动力学基础的必要性。

这种澄清沿着两个不同的方向进行。第一个方向由施温格(1966，1967)和温伯格(1967a，1968)开创，他们发展了一种对低能强相互作用的现象学研究进路，即所谓手征动力学，以使人们能够找到关于如何超越软 π 介子限制的线索。这一思想很简单。如果人们构造了一个手征不变的有效拉格朗日函数，并计算了最低阶的图，那么部分守恒轴矢流和等时对易关系及其结果将会重现。

尽管两者都采用了一种现象学进路，然而在施温格和温伯格之间仍然存在着重要的不同。在施温格的现象学源理论中，在最低阶中耦合项的使用是通过用数字(即非算符)表示的有效拉格朗日函数的真正本质来证明其合理性的。施温格还主张，只要对称性起源仍然模糊，现象学进路就是适合的。然而，对温伯格来说，仅在最低阶上进行计算的理由是为了重现部分守恒轴矢流—流代数方法的结果，环的效应已经通过形状因子 f_V 和 g_A 的出场得以说明。至于对称性方

[246] 面，与施温格相反，温伯格认为，当他还不知道如何从一个基本的拉格朗日函数推导出某种对称性时就在现象学层面上使用这个对称性，他是感觉不舒服的。 显然，即使当温伯格也从事发展现象学进路时，与施温格相比，他是更少现象学导向的。

历史上，现象学进路的重要性多半不在于它适用于各种低能介子过程，这经常可以从形式的(非动力学的)研究进路上推导出来，因为在其功能上现象学进路是有效场论的一个先驱，它在有效场论的发展中有启发性价值，而有效场论是从一个完全不同的语境中发展出来的(见 11.4 节)。 如果我们考虑后来的纲领的创始人之一又是温伯格这一事实，情形就特别是这样。

在澄清部分守恒轴矢流—流代数方法的动力学基础的研究上，第二个方向是发展了关于这一主题的一个根本性理解。 在这个方向上的工作持续了若干年，在这个领域的集体经验的累积最终导致了对于反常对称性破缺的随之而来的关注，这彻底改变了量子场论的概念基础和我们对于重正化的理解。 我们将在 8.7 节给出直至 1969 年的工作的一个说明。 揭示反常对称性破缺观念的丰富内涵的进一步发展将在 8.8 节和 11.4 节考察。

8.7 反常

在 20 世纪 60 年代后半期，对于定域场论的反常行为的深入研究，沿着三条密切相关的进路进行：(i)作为对流代数采用的等时对易关系的形式(独立于动力学细节)操作的一个回应，主要由约翰孙、洛、贝尔和其他人承担；(ii)作为对修正部分守恒轴矢流关系的必要性的一个回应，主要由韦尔特曼、萨瑟兰、贝尔、贾基夫、阿德勒和其他人作出；(iii)作为对流的重正化研究的一部分，主要由阿德勒、

巴丁、威尔逊和其他人进行。这些研究强有力地确立了定域场论某些反常行为的存在，给部分守恒轴矢流—流代数纲领采取的形式操作的有效性范围划定了界限，并使得微扰理论作为这一纲领的动力学基础变得更为清楚，因而结束了这一纲领作为一个自洽研究[46]框架的思想，为物理学家回到拥有各种精确和破缺的对称性的定域场论的框架铺平了道路，而在定域场论中，重正化是最迫切需要解决的问题。

反常对易式

[247]

对植根于正则等时对易关系基础之上的流代数的连贯性或有效性的严重怀疑，首先由约翰孙和洛(1966)提出。这一怀疑的根源可以追溯到后滕和今村(Goto and I. Imamura，1955)以及施温格(1959)，其时流代数还不存在。戈托、伊马穆拉和施温格知道，在一个有正度规的相对论性理论中，对于一个矢量或轴矢流来说，在涉及 δ 函数而不是正则 δ 函数的空间导数的 V_0 和 V_j 的对易式的真空期望中，一定有额外的项——后来被称为施温格项。

施温格的观测结论被约翰孙捡起。1961 年，约翰孙发表了一篇在历史上和概念上同等重要的关于蒂林模型的论文(1961)。约翰孙在那篇论文中，(i)解决了在同时发生的时空点上如何定义奇异场算符(例如，流)的乘积问题；(ii)证明了这些乘积由于其奇异本性和规则化的必要性，因而不满足正则等时对易关系。约翰孙的论文除了运用场具有动力学标度维的思想，为稍后威尔逊关于算符乘积展开的研究工作(1969，1970b，c；见下文)提供了一个出发点以外，他的论文还由于把施温格的观测结论从真空期望扩展到了 V_0 和 V_j 的对易式的算符(即非真空)结构，因而也变得具有历史重要性。施温格的观测结论只涉及导致常数减法的运动学，而约翰孙的论文则深深触及定域场论的动力学。

随着流代数的出现和盛行，等时对易式的非正则行为由于隐含着流代数的连贯性，因而成为深入研究的一个主题。[47]在1966年出现的各种研究中，约翰孙和洛的论文似乎是最富有成效和最有影响力的。[48]通过用一个可重正化的相互作用夸克模型，并借助于一个后来称为比约肯—约翰孙—洛限制(Bjorken-Johnson-Low limit，BJL limit)的技术设计——通过这一设计，等时对易式能与已知的格林函数的高能行为相关联——约翰孙和洛计算了，到介子—重子耦合中的最低非零阶为止，对于全部16个狄拉克协变式和一个任意的内部群，一个介子和真空(或一个介子)之间的对易式的矩阵元。通过这一全面考察，约翰孙和洛令人信服地证明了，在对易式的矩阵元与某些费恩曼图(诸如三角形图；见下文)相关的某些情形中，对易式无疑有有限的额外项。正如约翰孙和洛所指出的[49]，这些额外项存在的[248]原因如下。任何定域的相对论性理论，无论多么收敛，必定足够奇异以至于不得不引入规则化，这本质上是一种极限处理方法。当计算涉及一个比对数发散更厉害的积分时，这一积分将经受两个或多个不是合法可互换的限制[50]，这就有一个极限模糊性，额外项就是从对这个极限模糊性的一种精心处理中产生的。

受约翰孙和洛论文的激励，贝尔(1967a)举例说明了可解李模型中反常对易式的某些方面，特别强调了截止规则化。贝尔建议，当截止是有限值时，正则对易式对于零能量定理是正确的；只是对于涉及无穷大能量的求和定则来说，反常对易式才变得重要。

洛(1967)和贝尔都认为，对易式的反常行为将出现在一些高能求和定则中，然而只是以一种抽象的方式(约翰孙和洛)，或者是以一种非现实的模型(贝尔)出现。在1969年以前，对于反常对易子的讨论一直与实验相脱节。1969年，卡伦和格罗斯(1969)在比约肯(1966，

1969)提出的正则等时对易关系的基础上，发表了他们关于高能纵向电致截面的反常求和定则的推导，这立即受到了来自贾基夫和普雷帕拉塔(Jackiw and Preparata，1969)，同时还有阿德勒和吴大峻(1969)运用约翰孙—洛型反常对易式作为论据发起的挑战。

部分守恒轴矢流的修正

研究反常的另一条进路与实验有着更为密切的联系。作为历史事实，这直接受到在部分守恒轴矢流预测和介子衰变的观测结果之间引起争议的不一致的激励。1966年末，萨瑟兰发现，在 π 介子有等于零的四维动量($q=0$)的情形中，如果用部分守恒轴矢流和流代数的假定计算，$\eta\to 3\pi$ 衰变将被禁戒。[51]这直接与实验相矛盾。这里，部分守恒轴矢流以两种方式被使用。第一，轴矢流通过把它的发散当作内插的 π 介子场进行计算。第二，作为 q^2 的一个函数，计算的衰变幅从 $q^2=0$ 一直延拓到 $q^2=m_\pi^2$。由萨瑟兰提出的部分守恒轴矢流在实验上失效的一种可能解释是：由于一些未知的机制，这一延拓可能是非法的。这种未知机制如，与短程W介子相互作用相比，光子相互作用的长程性质会造成差异，这两种相互作用都被假定是这一衰变的原因。这一意见隐含着在第二种用法中修正部分守恒轴矢流。里亚兹丁和萨卡尔的论文(Riazuddin and Sarker，1968)，就是我们可以找到的沿着这一进路所作的研究。 [249]

韦尔特曼(1967)推广了萨瑟兰的观测结果；萨瑟兰的观测结果的大意是：任何包含一个 π^0 的过程，诸如 $\pi^0\to 2\gamma$ 和 $\omega\to\pi^0\gamma$，当用部分守恒轴矢流和流代数计算时，必定被禁戒于 π^0 零四维动量限制中。[52]根据这一观测结果，韦尔特曼作出结论：通常的部分守恒轴矢流(方程(14))是不正确的，必须增加包含 π 介子场的梯度的额外项，以便能够说明观测到的衰变。韦尔特曼提出的意见隐含着在第一种

用法中修正部分守恒轴矢流。沿着这一进路的还有阿诺威特、弗里德曼和纳特(Arnowitt, Friedman and Nath, 1968)，尽管他们所提出的额外项的形式不同于韦尔特曼提出的形式。

贝尔深入涉足关于流代数和部分守恒轴矢流的研究，韦尔特曼的观测结果给了他特别深刻的印象。[53]在大量的研究工作之后[54]，贝尔及其合作者贾基夫共同发表了一篇有趣的论文(1969)。在这篇论文中，他们声称“要用一个很简单的例子来证明，流代数计算共有的形式操作的不可靠性”。[55]他们也试图“提供一个方案，在这个方案中，对形式推理而言，事物被充分好地确定，从而可以进行明确的计算”，即这是一个“同时体现部分守恒轴矢流和规范不变性思想”的方案。贝尔和贾基夫调和这两种表面上矛盾的研究的方式是迷人的，也显露了他们的研究已触及当代理论物理学的概念预设有多深。

贝尔和贾基夫注意到，在 σ 模型中考虑过的 $\pi^0\rightarrow 2\gamma$ 衰变提出了一个疑难。一方面，关于韦尔特曼—萨瑟兰型的部分守恒轴矢流推理隐含着对于 $\pi^0\rightarrow 2\gamma$ 的不变幅，当用一个轴矢流发散的 π 介子场进行离质壳延拓时，$T(k^2)$ 消失了，即 $T(0)=0$。这个结论既与实验相冲突，又与施泰因贝尔格(Steinberger, 1949)关于为这一过程作出贡献的三角形曲线图的老的微扰计算相冲突，而这一微扰计算与实验符合得极好。[56]另一方面，在 σ 模型中同一过程的明确计算把部分守恒轴矢流固定在一个算符方程中，刚好给出与施泰因贝尔格的计算相同的结果。这是贝尔和贾基夫所面对的疑难。

为了揭示这一难题的起源，贝尔和贾基夫超越这一形式推理，考察了这一计算的细节。首先，他们发现，在 σ 模型计算中获得的非零 $T(0)$ 来源于一个表面项，当改变三角形曲线图的线性发散积分中的变量时，就获得了这一非零计算。这个额外项与部分守恒轴矢流

相一致，但破坏了规范不变性。因而这是在部分守恒轴矢流和规范不变性之间的一个冲突。其次，他们追随约翰孙和洛，主张：相关 [250] 三角形曲线图的积分的形式操作，尽管既遵守部分守恒轴矢流又遵守规范不变性，但是如果被积函数没有很好地收敛，它也可能失效。这导致表面项破坏规范不变性，并要求规则化。他们强调，传统的泡利—维拉里规则化虽然恢复了规范不变性，但却破坏了部分守恒轴矢流。部分守恒轴矢流和规范不变性之间的这种冲突再次出现。贝尔在他 1968 年 9 月 2 日写给阿德勒的信中评论说：

> 我们的第一个观测结论是，以传统方式解释的 σ 模型正好没有部分守恒轴矢流。这是这一难题的一个解决方案。[57]

然而，一个深深地植根于对称性崇拜之中的强有力的预设，使得贝尔和贾基夫不满足这一解决方案。这种对称性崇拜已为许多当代理论物理学家广泛拥有，并展示在我们上面刚刚提到的论文的第二个动机之中。贝尔和贾基夫试图用一种他们所关心的部分守恒轴矢流和规范不变性所共同规定的不同的方式来解决这一疑难。

与贝尔先前的论点(1967a, b)——这种论点认为，对易式中的反常是不相关的，它们能通过引入规则化而避免——相符合，贝尔和贾基夫提出了一种受部分守恒轴矢流规定的新的规则化。这里，关键的假定是调节子场(regulator field)的耦合常数 g_1 一定随着场的质量 m_1 以如下方式变化：$m_1/g_1 = m/g =$ 常数，即使当 m_1 趋于无穷且由真实场和辅助场造成的反常相互抵消时也是如此。因此，贝尔和贾基夫不是去修正部分守恒轴矢流，而是去修正 σ 模型中的规则化方案，以使得反常连同部分守恒轴矢流和规范不变性之间的冲突以及

形式推理和明确计算之间的差异，统统被消除。

不幸的是，贝尔和贾基夫修复的形式推理的连贯性经不起详细的审查。正如阿德勒所指出的(1969)，新的规则化将使得 σ 模型中的强相互作用不可重正化。[57a]阿德勒也指出，反常是那么深地植根于定域场论的理论结构之中，以至于不能被任何人为的操作去除。

尽管贝尔—贾基夫的论文缺乏有创见的动机，并且其解决方案也难以立足，然而它的确偶然发现了一些重要的事情。它的重要性在于两个方面：(i)它用可观测的 $\pi^0 \rightarrow 2\gamma$ 衰变表述在三角形曲线图的计算中出现的反常，因而在物理学家中间产生了对于反常的巨大兴趣，并把物理学家的注意力吸引到反常上来；(ii)它把反常与部分守恒轴矢流联系在一起，因而为理解定域场论的反常行为的本质指明了正确的方向。然而，若不是阿德勒(1969)以一种更为完备、更为优美的表达形式阐明这一反常情形的话，那么这篇论文的真正重要性不可能被及时地和适当地注意到。在回应贝尔和贾基夫关于部分守恒轴矢流疑难的解决方案时，阿德勒不是去修正 σ 模型以恢复部分守恒轴矢流，而是在传统的 σ 模型中尽力系统化和充分利用部分守恒轴矢流破缺，因而给出了 $\pi^0 \rightarrow 2\gamma$ 衰变的合理说明。然而，他在论文的附录中所作出的这一回应，只是他在论文正文中所报告的主要思想的一个应用(1969)。这些思想是在他关于旋量电动力学的研究中，独立于贝尔—贾基夫的论文而发展的，它们受重正化考虑的驱动，是在一种很不同于部分守恒轴矢流—流代数纲领的语境中实行的。[58]

[251]

流的重正化

阿德勒关于量子电动力学中轴向矢量顶点的重正化的动力学研究，是他作为一个活跃于部分守恒轴矢流—流代数纲领中的理论学家跨出的重要一步，这与通过纯粹的对称性考虑而不是诉诸动力学来析

取体现在顶点形状因子中的物理信息的纲领相背离。然而，在阿德勒的研究中，仍然容易探测到他对于对称性的残留迷恋。对阿德勒来说，这一研究是有趣的，因为它与无质量量子电动力学的 γ_5 不变性建立了关系。准确地说，因为这一研究是带着对它所具有的对称性的强烈兴趣进行的，所以在动力学计算和由对称性考虑规定的形式代数演算之间所发现的冲突，被有意识地充分利用，以揭示它对于我们理解由定域场论描述的系统的动力学对称性的重要性：当考虑由环状曲线图表征的辐射修正时，一些展示在拉格朗日函数中的对称性可能消失，而额外（“反常”）项可能会出现在相应的计算中。因此，对称性考虑的有效性范围由相关对称性破缺的出场或缺场来划界，这是由辐射修正的特殊过程所造成的。然而，在探究反常的含义之前，让我们先来更为切近地看一看阿德勒所采取的步骤。

在微扰研究中，阿德勒发现，除了定义费恩曼图的积分是线性发散或更严重的发散、因而积分变量的移位是不确定的情形之外，在量子电动力学中轴向矢量顶点一般满足通常的沃德恒等式。检测表明，量子电动力学中唯一令人烦恼的情形就是那些涉及三角形曲线图的情形。阿德勒发现，运用罗森伯格(Rosenberg)的三角形曲线图的明确表示(1963)，并运用通常的泡利—维拉里规则化来计算这种三角形曲线图时，如果这一计算与规范不变性相符合，那么轴向矢量沃德恒等式在三角形曲线图的情形中就会失效，一个额外项就会显现。[59] [252] 在阿德勒的表述中，轴向矢量沃德恒等式的失效是两个因素相结合的结果。第一个因素是三角形曲线图的一个特殊性质，即它的线性发散导致了对于曲线图的表达上的歧义。第二个因素是由规范不变性强加的限制。

在微扰理论中，沃德恒等式由基本拉格朗日函数的不变性（或部

分不变性)及其表示所规定。 因此，轴向矢量沃德恒等式的失效已经隐含着一种对称性破缺。 然而，阿德勒更为清楚地说明了在它的反常发现和对称性破缺之间的联系。 阿德勒注意到轴向沃德恒等式中破缺的出现，不但适合于基本的三角形曲线图，而且适合于用基本的三角形曲线图作为亚曲线图的任何曲线图，他表明，在一般情形中的破缺能通过取代通常的轴矢流发散而加以简明描述：

$$\partial j_{\mu}^{5}(x)/\partial x_{\mu} = 2\mathrm{i}m_{0}j^{5}(x) \tag{24}$$

$(j_{\mu}^{5}(x)^{-} = \psi(x)\ \gamma_{\mu}\gamma_{5}\psi(x)$，$j^{5}(x)^{-} = \psi(x)\ \gamma_{5}\psi(x))$ 通过

$$\partial j_{\mu}^{5}(x)/\partial x_{\mu} = 2\mathrm{i}m_{0}j^{5}(x) + a_{0}/4\pi{:}F^{\xi\sigma}(x)F^{\tau\rho}{:}\ \varepsilon_{\xi\sigma\tau\rho}, \tag{25}$$

阿德勒把这看作是他的主要结果。 通过把这个结果应用于无质量的量子电动力学情形，阿德勒证明了轴向矢量三角形曲线图的出现，提供了一个适合于破坏无质量量子电动力学的拉格朗日函数 γ_5 不变性的特殊机制。 在这种情形中 γ_5 不变性被破坏，完全是因为与 γ_5 变换相联系的轴矢流由于额外项 $a_{0}/4\pi{:}F^{\xi\sigma}(x)F^{\tau\rho}{:}\ \varepsilon_{\xi\sigma\tau\rho}$ 而不再守恒之故。[60]

阿德勒将导致方程(25)的相同思想应用于 σ 模型的贝尔—贾基夫版本，他发现相同的额外项也出现在部分守恒轴矢流关系式的右边。这需要部分守恒轴矢流必须在电磁相互作用出现的情况下以一种定义明确的方式加以修正。 阿德勒修正的部分守恒轴矢流关系式给出了有关 $\pi^{0}\rightarrow 2\gamma$ 衰变($\tau^{-1} = 9.7\ \mathrm{eV}$)的一个预测[61]，它不同于由萨瑟兰、韦尔特曼以及贝尔和贾基夫给出的那些预测，它与实验值 $\tau_{\exp}^{-1} = (7.35 \pm 1.5)\mathrm{eV}$ 符合得很好。

这里，一个重要然而却不明显的假设是：反常是一个精确的结果，它对所有阶都有效，也不为高阶辐射修正重正化。 在阿德勒的

分析中，这个假设自然地源于这样一个事实：对于基本三角形曲线图 [253]
的辐射修正，总是涉及多于三个顶点的轴向矢量环，这在量子电动力学的情形中是处在对数发散的最坏情况下，因而满足法向轴矢沃德恒等式。 作为对贾基夫和约翰孙(1969)提出的这一假设的不同意见的反应，阿德勒和巴丁(1969)设计出说明这一假设的论据的细节。 这导致了一个非重正化定理，这一定理对于反常消除条件是至关重要的；也对由韦斯和祖米诺(Julius Wess and Bruno Zumino， 1971)以及特霍夫特(1976a)提出的连贯性条件是至关重要的。[62]

关于轴向矢量顶点的重正化，阿德勒注意到，包括额外项在内的轴向矢量发散不是倍增地可重正化，这影响深远地使得轴向矢量顶点的重正化变得复杂了，虽然通过为不同的费米子种类安排反常以致相消，能够避免出现这些复杂因素。[62a]

阿德勒对贝尔和贾基夫的新规则化的批评，据说有助于去除反常，运用这种批评，与定域场论的反常行为相关联的概念状况得以大大澄清。 这种批评令人信服地表明，如果不破坏规范不变性、幺正性或可重正化性，就不能消除三角形曲线图反常。[63] 以此为据，三角形反常不久就成了已确立的知识的组成部分。

首先在阿贝尔理论中被很好地确立的阿德勒—贝尔—贾基夫反常，不久就为格斯坦和贾基夫(Gerstein and Jackiw，1969)以及巴丁(1969)扩展到非阿贝尔情形。 这一扩展与服从可重正化性约束的标准模型的构造有直接关系(见 10.3 节)。 格斯坦和贾基夫认为，在 SU(3) × SU(3) 耦合情形中，除了有一个反常的 VVA 三角形曲线图以外，唯一的曲线图是 AAA 曲线图，而巴丁则声称有一个更一般的 SU(3) 反常结构。 对中性流而言，它与格斯坦和贾基夫的结论相一致；对带电流来说，反常也在 VVVA 和 VAAA 方格曲线图和

VVVVA、VVAAA 和 AAAAA 五角形曲线图中出现。

短程算符乘积展开

关于流的重正化效应，几乎同时也受到威尔逊(1969)的攻击。[64] 威尔逊不涉及拉格朗日函数短程算符乘积展开(operator product expansions，OPE)的框架，以及他对标度不变性而不是 γ_5 不变性的关心，非常不同于阿德勒和贝尔—贾基夫。然而，威尔逊的分析为理解重正化和对称性破缺之间的关系指出了相同的方向，对于后来的概念发展有相似甚至是更重要的意义。

[254]

自约翰孙和洛的工作以来，流代数在处理流的短程行为中的不足已为人熟知，作为对这一不足的回应，威尔逊试图发展一种新的表达方式，即算符乘积展开，这种表达方式能给出强相互作用中流的短程行为的一幅更详细的图景：

$$A(x)B(x) = \sum_n C_n(x-y)O_n(x) \tag{26}$$

这里，$A(x)$，$B(x)$和 $O_n(x)$可以是任何定域场算符(基本场、流、能量张量，等等)；$C_n(x-y)$——包括$(x-y)$的幂和$(x-y)^2$ 的对数，还可能在光锥上有奇点——包含了所有关于流的短程行为的物理信息。

算符乘积展开起源于微扰理论中对重正化的详细研究。[65] 然而，威尔逊现在是在破缺的标度不变性的新基础上发展它[66]，他利用破缺的标度不变性来确定 $C_n(x-y)$ 的奇异性结构。其出发点是约翰孙关于蒂林模型的工作，约翰孙的工作已表明(i)正则对易式受到重正化效应的破坏，(ii)场的标度元随着耦合常数而连续变化。威尔逊进行了推广，使得这两个观测结果在他关于强相互作用的算符乘积展开表述的情形中也得以保持。威尔逊在论文中竭力主张的主要观点之

一是：从流的非相互作用数值中导出的流的标度元的存在，令人信服地表明重正化效应造成了标度不变性的破缺。这一意见也获得了约翰孙和洛的观测结果的支持；威尔逊也在论文中要求人们注意，正则等时对易关系将受到非不变相互作用的破坏。

威尔逊的分析似乎隐含着这样的观点：强相互作用的标度不变性，既不被拉格朗日函数中的对称性破缺项（在威尔逊的方案中没有这样的拉格朗日函数）所破坏，也不被一个可变真空所破坏，而是如同 γ_5 不变性的反常破缺一样，只被在重正化程序中引入的一些可变相互作用所破坏。这一观点不久就由威尔逊本人认真地作了考察（1970a，b，c；1971a，b，c；1972），也为卡伦（1970），西曼齐克（1970），卡伦、科尔曼和贾基夫（1970），科尔曼和贾基夫（1971），以及其他人所考察。这一考察直接导致重正化群思想的复兴。

正如我们在下一节中将要看到的，威尔逊对于标度不变性的反常破缺的态度与其他人很不相同。这一态度似乎发端于他1969年的论文。虽然承认反映在流的标度维（scale dimesion）变化中的标度反常存在，但是威尔逊坚持所有反常都能被吸收进流的反常维（anomalous [255] dimension）中，因此标度不变性将在渐近意义上保持，即标度律（scale law）仍然成立，虽然仅仅对带电标度维的流成立。

威尔逊的态度似乎可归因于：(i) 部分是盖尔曼和洛关于量子电动力学中裸电荷的标度律的工作（1954）对他的工作的影响；(ii) 部分是菲舍尔、维多姆和卡达诺夫（Fisher，Widom and Kadanoff）临界现象上的标度假设对于他的工作的影响；(iii) 部分是他企图使自己脱离微扰理论的雄心。前两点使他相信标度律的存在，或者用后来的术语说，使他相信重正化群的不动点的存在；最后一点使得他不能用一种完全动力学的方式来分析标度反常。标度反常的这一方面和其他

方面将是下一节的主题。

8.8 重正化群

20 世纪 70 年代初，对于重正化群方程的复兴和重新表述来说，概念基础是标度反常(scale anomaly)的思想，或标度不变性的反常破缺。然而，历史上，在量子场论的框架中，物理参量的标度依赖性思想，比标度不变性的思想出现得早，并为正确理解标度不变性提供了语境。

物理参量的标度依赖性思想的早期(甚至是第一次)出现，我们可以追溯到戴森关于平滑相互作用表象(smoothed interaction representation)的工作(1951)。在这一表象中，戴森试图把相互作用的低频部分和高频部分分开，除了产生重正化效应之外，高频部分被认为是无效的。为了实现这一目标，戴森采用了绝热假说的指导方针，借助于平滑变化的参量 g，为电子定义了一个平滑变化的电荷和一个平滑变化的相互作用。然后他指出，当 g 变化时，在依赖于 g 的相互作用的定义中必须作出某些修正，以使由依赖于 g 的电荷的改变造成的这一效应能得到补偿。

与戴森的思想相符合，朗道及其合作者在一系列有影响的论文(1954a，b，c，d；1956)中，发展了一个模糊化相互作用(smeared out interaction)的相似概念。依照这个概念，相互作用的量值不应该被看作一个常量，而应该被看作相互作用半径的函数，当动量超过临界值 $P \sim 1/a$(其中 a 是相互作用的范围)时，相互作用的大小必定迅速下降。[256] 相应地，电子电荷必须被看作一个至今未知的相互作用半径的函数。朗道借助这个概念，研究了量子电动力学的短程行为，并获得了一些重要的结果，这在开始的几节中已提及。戴森和朗道都

持有这样的思想：与电子电荷相对应的参量是标度依赖的。此外，戴森暗示——虽然只是含蓄地暗示——量子电动力学的物理学应该是标度独立的。朗道则更为明确地提出，量子电动力学的相互作用可能是渐近标度不变的。

“重正化群”(renormalization group)这一术语首先出现在斯蒂克尔伯格和彼得曼写的一篇论文(1953)中。在那篇论文中，他们注意到，虽然在重正化程序中引入的相反项的无穷大部分，是由消除发散的要求决定的，但是有限的部分是可改变的，它依赖于相减点的任意选择。然而，这种任意性在物理上是不相关的，因为不同的选择只是导致理论的不同参量化。他们评论说，能够定义一个变换群，这个变换群与该理论的不同参量相关，他们称之为“重正化群”。他们也指出了引入一个无穷大算符和构造一个微分方程的可能性。

一年后，盖尔曼和洛(1954)在对量子电动力学短程行为的研究中，更富有成果地开拓了重正化不变性。首先，他们强调，被测电荷 e 是量子电动力学动量非常低时行为的一个特性，并且 e 能被参量族 e_λ 中的任何一个成员取代，它与任意动量标度 λ 上的量子电动力学行为相关联。当 $\lambda \to 0$ 时，e_λ 变为被测电荷e，当 $\lambda \to \infty$ 时，e_λ 变为裸电荷 e_0。其次，他们发现，由于可重正化性，e_λ^2 服从方程 $\lambda^2 \mathrm{d}e_\lambda^2/\mathrm{d}\lambda^2 = \psi(e_\lambda^2, m^2/\lambda^2)$。当 $\lambda \to \infty$ 时，重正化群函数 ψ 变为只是e_λ^2 的一个函数，因而为 e_λ^2 确立了一个标度律。第三，他们指出，作为这个方程的一个结果，裸电荷 e_0 一定有一个不依赖于被测电荷 e 值的不变值；这就是所谓的盖尔曼—洛裸电荷本征值条件。

在斯蒂克尔伯格和彼得曼以及盖尔曼和洛的工作中，戴森的变化参量 g 和朗道的相互作用范围，被进一步指定为可调整的重正化标度或相减点。与盖尔曼和洛一致，参量的标度依赖特征和不同重正化

标度上的参量之间的联系，按照重正化群变换详加描述，并且此物理学的标度独立特征在重正化群方程中得以具体化。然而，这些苦心经营的结果直到晚近的 20 世纪 60 年代末和 70 年代早期才受到重视。[257] 其时，主要是通过威尔逊的研究，由于量子场论和统计物理学之间富有成效的交互作用，人们对于标度不变性思想和重正化群方程才有了更深入的理解。

理论的标度不变性的思想更为复杂，它非常不同于理论的物理学标度独立性的思想(由戴森模糊地提出)，或理论的物理学重正化标度独立性的思想(由重正化群方程表达)。理论的标度不变性指的是它在标度变换群下的不变性。后者只为动力学变量(场)下定义，不为量纲参量(诸如质量)下定义。否则，一个标度变换将导致一个不同的物理理论。虽然理论的物理学应该独立于重正化标度的选择，但是，如果有任何量纲参量，那么该理论就不可能是标度不变的。

在盖尔曼和洛关于量子电动力学的短程行为的处理中，当电荷按照它在很大距离上的值被重正化时，这个理论不是标度不变的。由于电子的质量可以忽略，并且似乎没有其他量纲参量出现在这一理论中，因此在这种情况下人们可以期望获得标度不变性。标度不变性意料之外的失效，完全归咎于电荷重正化的必要性：当电子质量趋于零时，有一个奇点。然而，当电荷在一个相关能量标度上，通过引入一个可调整的重正化标度，以有效地抑制不相关的低能自由度而被重正化时，就出现了一个渐近标度不变性。盖尔曼和洛用关于有效电荷的一个标度律来表达这个“渐近标度不变性”。他们也把它当作裸电荷的本征值条件，意思是指对裸电荷有一个不依赖于实测电荷值的不变值。

虽然约翰孙在 20 世纪 60 年代初期就提出，蒂林模型可能是标度

不变的，但是直到20世纪60年代中期，由于统计物理学的发展，才有了关于标度不变性本质理解的真正进展。此外，在流代数研究中和量子场算符乘积的短程展开中所发现的场论反常，也促进了这一领域的研究。

这里，量子场论和统计力学之间的交互作用——这在威尔逊的思想形成中可轻易察觉出——在这一发展中起了重要的作用。概念上，这一交互作用很有趣，但是也很复杂。1965年，维多姆在论文(1965a，b)中为靠近临界点的态的方程提出了一个标度律，就临界指数之间的关系而言，这推广了由埃萨姆和菲舍尔(J. W. Essam and M. E. Fisher，1963)以及菲舍尔(1964)早期获得的结果。威尔逊对维多姆的工作感到困惑，因为后者的工作缺乏理论支持。威尔逊熟悉盖尔曼和洛的工作。而且，当通过解决和消除这一问题的动量标度而致力于发展格场理论(lattice field theory)时，他刚好已为重正化群分析找到了一个自然的基础(1965)。威尔逊意识到应该把盖尔曼和洛的思想应用到临界现象中去。一年后，卡达诺夫(1966)把临界点变成在标度依赖的参量变换中的一个不动点，他从这一思想出发推导了维多姆的标度律，这本质上是把重正化群变换具体化了。威尔逊迅速吸收了卡达诺夫的思想，把它合并进他关于场论和临界现象的思考之中，拓展了破缺标度不变性的概念。[67]

[258]

威尔逊在1964年就算符乘积展开也做了一些开创性的工作(未发表)，但是，在强耦合领域，算符乘积展开失效了。在思考了维多姆和卡达诺夫标度理论应用到量子场论时的含义，并且研究了约翰孙关于蒂林模型的标度不变性的意见的结果(1961)和麦克(Mack)关于强相互作用的短程标度不变性的意见(1968)之后，威尔逊重新表述了算符乘积展开理论，并把它建立在标度不变性的新思想的基础之上

(1969)。他发现，如果把场算符的标度维——由正则对易关系的标度不变要求来定义——看作新的自由度，那么量子场论可能在短程上是标度不变的。[68]这些标度维能被场之间的相互作用改变并且能获得反常值[69]，这些反常值对应于临界现象中的非平凡展开。

对量子场论进行根本性改造的最重要影响来自统计物理学的巨大进展，这些进展可用威尔逊强调的两个概念加以总结：(i)定域理论中的统计连续统极限(statistical continuum limit)，以及(ii)重正化群变换的不动点。

就第一个概念而言，威尔逊注意到，由统计物理学描述的系统和由量子场论描述的系统，具体表达了不同的标度。如果一个连续变量的函数，例如在时空上定义的电场的函数，本身独立于变量并假设其形成一个连续统，以使函数积分和导数能定义，那么人们就能定义[259] 一个统计连续统极限，它由一个特征标度的缺失来刻画。这意味着在所有标度上的涨落相互耦合，并对过程所作的贡献相等。在量子电动力学的计算中，这典型地导致对数发散。因此，对于这些系统的研究，重正化是必要的。“统计连续统极限”这一概念在威尔逊关于量子场论、特别是关于重正化的思想中，占据着中心的位置；这反映在他如下的断言(1975)中：“标准重正化程序的最坏特征是，它没有给出关于统计连续统极限的物理洞见。”

第二个概念更为复杂。正如我们已注意到的，一个刻画不同重正化标度上的物理学的变化参量，反映了重正化效应的标度依赖性。这一参量的值通过重正化群变换相互关联，这些重正化群变换由重正化群方程描述。从20世纪60年代后期起，人们开始认识到，由于重正化的必要性，任何量子场论的标度不变性总是不可避免地被反常所破坏。这个论断的一个令人信服的论据是基于“量纲嬗变”

(dimensional transmutation)的概念，对此具体描述如下。

一个理论的标度不变性等价于在这个理论中标度流的守恒。为了定义标度流，就要求有一个重正化程序，因为作为同一点上两个算符的乘积，标度流隐含有一个紫外奇点。然而，即使是在一个没有任何量纲参量的理论中，为了避免红外发散和定义耦合常数，在重正化时引入一个量纲参量作为相减点仍然是必要的。引入一个量纲参量的这一必要性，科尔曼和温伯格在他们的论文(1973)中称之为"量纲嬗变"，它打破了这一理论的标度不变性。确切地讲，由于量纲嬗变，在任何重正化理论中的标度不变性不可避免地会被反常所破坏，虽然这一破缺效应能通过重正化群方程来处理。

在统计物理学中，重正化群方法影响了在不同标度层次上的物理学之间的联系。通过去除不相关的短程关联，并且通过确定稳定的红外不动点，重正化群方法使得各种描述[如基本激发(准粒子)的描述和集体粒子(如光子、等离子体、自旋波)的描述]的概念一致性，各种临界行为的普遍性的解释，以及序参量和关键分量的计算，成为可能。在量子场论中，相同的方法能被用来去除不相关的低能自由度，并找到稳定的紫外不动点。在两种情形中，正如温伯格(1983)已指出的，这种方法的实质集中在一个特殊问题的相关自由度上[70]，[260]其目标是找到重正化群方程的不动点解。

根据威尔逊，量子场论中的不动点只是对量子电动力学中裸电荷的盖尔曼—洛本征值条件的一个推广。[71]在不动点上，标度律，或者是在盖尔曼—洛—威尔逊的意义上，或者是在比约肯的意义上，保持有效，并且这个理论是渐近标度不变的。这一标度不变性在非不动点上破缺，这种破缺能用重正化群方程描述。因此，运用更为精致的标度论据，盖尔曼和洛的原初思想的含义就变得更为清楚了。也

就是说，以一种定量的方式，而不是如戴森或朗道所建议的定性研究，即遵循由标度不变性的反常破缺引起的、理论的有效参量随着能量标度的改变而变化，重正化群方程能被用来研究各种能量标度上，特别是在很高能量标度上的场论的性质。

显然，如果一个给定场论的重正化群方程拥有一个稳定的紫外不动点解，那么这个理论的高能行为不会造成任何麻烦；根据温伯格，这个理论可以叫做“渐近安全理论”(1978)。如果一个渐近安全理论拥有的不动点是高斯不动点，那它可能是一个可重正化理论。[72]然而，温伯格指出“渐近安全”的概念比可重正化性的概念更为普遍，因此能用“渐近安全”的概念解释甚至取代可重正化性的概念，并且他用一个五维标度理论的具体例子来支持他的观点(同上)。实际的情形可能是，如果理论是与威尔逊—菲舍尔不动点相联系的，那么在通常的意义上理论是渐近安全的却不是可重正化的。

在这节中描述的概念发展可以总结如下。在有许多相互耦合的标度但却没有一个特征标度(如由量子场论描述的那些标度)的系统中，由于重正化的必要性，标度不变性总是反常破缺。这一破缺自我显现在算符乘积展开框架中场的反常标度维中，或者出现在由重正化方程具体规定的不同重正化标度的参量的变化之中。如果这些方程没有不动点解，那么这一理论不是渐近标度不变的，因而严格地说不是可重正化的。[73]另一方面，如果方程拥有不动点解，那么理论就是渐近标度不变的和渐近安全的。如果不动点是高斯的，那么理论是可重正化的。但是，可能有一些渐近安全理论，如果它们拥有的不动点是威尔逊—菲舍尔不动点，那么它们是不可重正化的。随着渐近安全是更为基本的指导原则这样一种思想的出现(这是重正化群方法的结果之一)，可重正化性原则的基本性开始受到严重的挑战(见

[261]

11.4节）。

8.9 来回摆动

在20世纪40年代末、50年代初，经历了对于量子电动力学的重正化微扰表述引人注目的成功的短暂沉迷之后，物理学家逐渐远离了量子场论的理论框架，这主要是因为量子场论在表达电磁学领域以外的各种问题上失效了。非常有趣的是，我们注意到这种变化始于20世纪20年代末海森伯和泡利的量子场论，然后又摇摆到20世纪30年代末、40年代初惠勒和海森伯的S矩阵理论。但是，更为有趣的事实是：20世纪40年代末物理学家皈依量子场论的重正化微扰表述，却是打算通过再次摆回到量子场论的框架来实现，虽然这次量子场论采取了规范不变理论的形式。

正如我们在先前的部分中所表明的，量子场论在20世纪50年代处于深刻的危机之中。除了极少数在公理化场论方面工作或在可重正化非阿贝尔规范理论方面思考的物理学家以外，主流物理学家为两种趋势所吸引。在强相互作用领域，主导思想是关于色散关系、雷杰极点和靴袢假说的思想，它们由丘及其同事发展为一个反场论框架，即解析S矩阵理论。在弱电磁过程和强子电磁过程的领域中，以及强相互作用重正化的领域中，形式推理和代数计算取代了场论的研究。在这两种情形中，场论框架即使没有被彻底放弃，至少也是受怀疑的。

但是，从20世纪60年代中期起，情形开始改变。在强相互作用领域，随着20世纪60年代在j复平面上越来越复杂的S矩阵的奇异结构的发现[74]，S矩阵理论逐渐衰退。那些依然从事S矩阵理论的丘的追随者，变成了一群从事远离粒子物理学研究主流的孤立的专

家。相比之下，主要是由于1969年在斯坦福直线加速器中心进行的实验发现了在深度非弹性散射中的标度律，因此夸克—部分子模型(quark-parton model)，以及在场论框架中更为复杂的量子色动力学(quantum chromodynamics)，被接受为强子物理学的基本框架。

[262] 在弱过程和电磁过程领域中，经由反常发现的刺激，形式操作再次被它们的动力学基础的场论研究所取代。而且，在20世纪60年代后期，在非阿贝尔规范理论的框架内，人们提出了用一种统一的方式来描述弱过程和电磁过程的模型。随着韦尔特曼和特霍夫特对非阿贝尔规范理论的可重正化性的证实，到1972年，粒子物理学家得到了弱力和电磁力的一个自洽的统一场论。接下来要做的事是获得由这一理论所预言的中性流的存在的实验证据。1973年这一证据出现了。

强相互作用中标度律的实验发现和弱相互作用中中性流的实验发现，使得预言或解释这些发现的非阿贝尔规范理论在20世纪70年代特别流行。但是，改变人们对于规范理论(通常对于量子场论)的流行看法的是其可重正化性的证据。在流行看法上的改变是那么激进，以至于朗道反对量子场论的颠覆性论证(虽然它不可能挑战渐近自由的杨—米尔斯理论，即重正化电荷的可能消失)，完全被忘记了。

这在一个更为基本的层次上证明，在近40年的理论发展中，各种量子场论，除了一些非现实的模型以外，只能通过引入规范不变性的概念来实现重正化。由于这个原因，量子场纲领作为一个连贯和成功的研究纲领，它的进一步发展几乎完全是在一个规范理论的框架之内。

注释

1. 注意，我在这段引文中提到的场算符处于海森伯绘景中，这与在相互作用绘景中有一个不同的物理解释。

2. 对于早期的W介子理论，见Kemmer(1938)和Klein(1938)。他们在解释π衰变$\pi \rightarrow \mu + \nu$中遇到了困难。W介子理论的现代版本由施温格(1957)和格拉肖(1961)在他们统一电磁学和弱相互作用的早期尝试中提出。但是这只在夸克模型的框架中才有意义(见Weinberg，1967b；Salam，1968)。

3. 然而，这是有争议的。见Weinberg(1983)，Gell-Mann(1987)。

4. 一些公理化场论学家后来以构造场论学家著称。怀特曼(Wightman，1978)把构造量子场论当作公理化场论的产物，但两者之间有一些不同。公理化场论关心的是量子场的广义理论，而构造场论从特定的拉格朗日模型出发，构造满足前者要求的解。关于公理化场论的早期发展，见Jost(1965)和Streater and Wightman(1964)；关于构造场论，见Velo and Wightman(1973)。

5. 这个结果通常不被接受。因此洛(1988)在过去和现在都不相信朗道的证据，因为 [263] 它基于微扰方法中一个重要的对数近似，他觉得它作为一个证据还不够严密。波尔金霍尼(1989)持相似的立场。盖尔曼(1987)也持这样一种立场，他1956年访问莫斯科时，曾与朗道就这一点展开争论。盖尔曼也评论道："用今天的语言来说，朗道大概想要说非渐近自由场论是不令人满意的。"(同上)

6. 在量子场论框架中，关于这一主题的最早研究出现在1957—1958年(Nishijima，1957，1958；Zimmermann，1958；Haag，1958)。其结果是，就散射理论而言，复合粒子和基本粒子之间没有任何不同。因此，在来自S矩阵理论方面的压力日益增加的情况下，出现了把"Z=0"猜测为粒子合成的一个判据，这首先由瓦尔和茹韦(J. C. Houard and B. Jouvet，1960)以及沃恩、阿龙和阿马多(M. T. Vaughn，R. Aaron and R. D. Amado，1961)提出，并证实了李模型(Lee，1954)。这个主题很流行，但只讨论了一些简化模型场论。进一步的参考能在温伯格(1965a)的论文中找到。

7. 克拉默斯(1944)提出了解析性的重要性；克勒尼希(Kronig，1946)把解析性与因果性联系起来。

8. 关于历史评论，见Cushing(1986，1990)。

9. 根据盖尔曼在罗切斯特会议上的演讲来判断，他那时似乎不是一个鲜明的场论学家。他用这样的期待结束演讲："可能这种方法[色散纲领]将阐明如何构造一个不同[于旧场论]的新理论，新理论将有可能解释场论无论如何也不可能解释的高能现象。"(1956)

盖尔曼的罗切斯特演讲对于S矩阵理论的产生是一个强有力的推动，这一论断为以下事实所证实：(i)曼德尔施塔姆的双重色散关系(1958)——这通常被认为是S矩阵理论的产生中具有决定性的一步——始于盖尔曼的色散关系可能取代场论的猜测；(ii)丘在1958年的日内瓦会议上引用了盖尔曼的罗切斯特评论；(iii)除了对S矩阵理论的概念发展的各种贡献外，正如我们在这段引文中所提到的，盖尔曼也是承认雷杰工作的重要性并支持靴袢假说的第一批物理学家之一。正如丘在1962年所评论的，盖尔曼"多年来对S矩阵理论和我施加了重要的积极影响；他的热情和过去几个月的敏锐观察已显著地加速了这些事件的进程，也刺激了我个人的兴奋感"(1962a)。

然而，盖尔曼后来关于S矩阵理论地位的立场很不同于丘。对于丘来说，S矩阵理论是理解强相互作用的唯一可能的框架，因为量子场论在这方面毫无希望。然而，对于盖尔曼来说，基本框架是量子场论，因为S矩阵理论的所有原理和属性实际上是从量子场论抽象出来或为量子场论所提示的，S矩阵理论本身只能被看作一种详细说明场论的方式，或是一个在质壳上的场论。盖尔曼对于S矩阵理论和量子场论之间关系的立场是处在费恩曼和戴森，以及莱曼、西曼齐克和齐默尔曼(1955)的传统中；在这个传统中，S矩阵被看作可以从场论导出，或至少与场论框架不是不相容的。但是对于丘来说，量子场论和S矩阵理论之间的冲突是不可调和的，因为由量子场论假设的基本粒子的观念导致了与解析

性和靴袢观念的冲突。

10．在关于分波振幅的色散关系中，卡斯蒂列霍、达利茨和戴森（L. Castillejo，R. H. Dalitz and Dyson，CDD）在此振幅中发现了一些额外的固定极点，因此它不能完全由输入分波振幅决定（1956）。CDD 极点通常与输入基本粒子相对应，但是，正如在 20 世纪 60 年代后期发现的，它们也可能出现于非弹性截止（见 Collins，1977）。

11．关于色散关系和解析 S 矩阵理论之间的这个关系的讨论，见 Chew（1989）。

12．Blankenbecler，Cook and Goldberger（1962）.

[264] 13．Charap and Fubini（1959）.

14．Chew and Frautschi（1961a，b）.

15．在这种情形中，s 表示在 t 信道中的动量传播子平方，t 表示能量平方。

16．关于 S 矩阵理论的本体论承诺的讨论，见 Capra（1979）。

17．盖尔曼对于基本粒子雷杰化的最初动机的论述，见他的论文（1989）；对于雷杰化纲领的历史研究，见 Cao（1991）。

18．见 Chew（1989）和 Capra（1985）。

19．Landau（1959）. 库特科斯基（R. E. Cutkosky）在 1960 年的论文中给出了一个包括幺正性条件的进一步的规则。

20．因此，在朗道图和费恩曼图之间有一种一对多的对应。

21．这部分是由于在 S 矩阵理论中重点从客体（粒子）转移到过程（散射或反作用）。

22．Stapp（1962a，b；1965；1968）；Chandler（1968；with Stapp，1969）；Coster and Stapp（1969；1970a，b）；Iagolnitzer and Stapp（1969）；and Olive（1964）.

23．见 Chew（1989）和 Capra（1985）。

24．丘在美国物理学会的讲演（1962 年 1 月）的题目是"没有基本粒子的强相互作用的 S 矩阵理论"（1962b）；他在日内瓦会议（1962 年 7 月）上讲演的题目是"没有基本粒子的强相互作用理论"（1962c）。

25．费恩曼在 1961 年关于基本粒子的普罗旺斯地区艾克斯会议上的演讲表达了这个原理。见洛论文（1962）中的脚注 2。费恩曼的这一思想被丘和弗劳奇（1961a）、萨拉姆（1962a）以及 20 世纪 60 年代早期的许多其他物理学家广泛引用。

26．盖尔曼的评论在（1987，1989）论文中，洛的评论在他（1988）论文中。

27．Chew and Mandelstam（1960），以及 Zachariasen and Zemach（1962）。

28．见 Cao（1991）。

29．Capps（1963）；Cutkosky（1963a，b，c）；Abers，Zachariasen and Zemach（1963）；以及 Blankenbecler，Coon and Roy（1967）。

30．两年以前，布兰肯贝克勒、库克和戈德伯格（Blankenbecler，Cook and Goldberger，1962）已试图把 S 矩阵理论应用于光子。关于这点的更多讨论能在 Cao（1991）找到。

31．关于从纯 S 矩阵理论证明因果场算符的存在，见 Lehmann，Symanzik and Zimmerman（1957）；关于把量子场论还原到一个双重模型的极限情形，见 Neveu and Sherk（1972）；关于把超引力理论和超杨—米尔斯理论作为超弦理论的极限情形，见 Green，Schwarz and Witten（1987）。

32．关于从场算符推导 S 矩阵元，见 Lehmann，Symanzik and Zimmerman（1955）；关于从因果性要求推导量子格林函数的色散表象，见 Nambu（1955）。

33．对于雷杰化纲领的初步研究，是量子场论和 S 矩阵理论之间概念相互影响的一个产物，能在 Cao（1991）的论文中找到。

34．PCAC 是"部分守恒轴矢流"的首字母缩写词。

35．这个延伸预设了弱流的轴向矢量部分也是守恒的。

36．$f_v(0)$ 由 $(2\pi)^3\ \langle p' \mid V_\mu(0) \mid p\rangle = iu(0)\gamma_4 f_v(0)u(0)\delta_{\mu 4}$ 定义，当 $p = p' \to 0$ 时，这是方程（2）的极限情形。

37．CVC 假说隐含着对强相互作用 $f_v(0)$ 不可重正化，这一点的简单证明可以在 Coleman and Mandula（1967）中找到。

38．定义 $T_+ = \int dr(J^v_{01}(r, 0) + iJ^v_{02}(r, 0))$。根据同位旋对易关系，我们有 $J^v_{\mu+} =$

$[J^{v}_{\mu 3}, T_{+}]$，因此

$$\langle p' |_{p} J^{v}_{\mu +} | p\rangle_{N} = \langle p' |_{p}[J^{v}_{\mu 3}, T_{+}] | p\rangle_{N} = \langle p' |_{p} J^{v}_{\mu 3} | p\rangle_{p} - \langle p' |_{N} J^{v}_{\mu 3} | p\rangle_{N}$$
$$= \langle p' |_{p} j^{em}_{\mu} | p\rangle_{p} - \langle p' |_{N} j^{em}_{\mu} | p\rangle_{N}$$

这里，T_{+} 的属性是作为一个同位旋提升算符，并运用了 $j^{em}_{\mu} = j^{s}_{\mu} + j^{v}_{\mu}(= J^{v}_{\mu 3})$。如果我们定义 $F_1(q^2)^{+} = f_v(q^2)$，$F_2(q^2)^{+} = f_m(q^2)$，那么方程(6)就是从方程(1)和方程(2)得出的。[265]

39. 注意，与规范 SU(3)对称性(即色对称性)不同，SU(3)对称性(及其亚对称性SU(2))在 20 世纪 60 年代初期很流行，它是一种整体对称性，后来被称做味对称性，是质量简并性的显示，先是三个坂田子[坂田昌一在(1956)论文中提出的核子的组分](或者在SU(2)情形中的质子和中子)，后来是夸克。在无质量的限制中，SU(3)或 SU(2)味对称性变成 $SU(3)_L \times SU(3)_R(SU(2)_L \times SU(2)_R)$手征对称性。

40. 我们将不给出在这个对易关系的定义中所涉及的所有复杂因素。关于这些细节，见 Gell-Mann(1962a，1964a，b)。

41. 在方程(3)的极限情形中，$g_A(0)$以相似于 $f_v(0)$的方式来定义(见注释 35)。

42. 这个束缚态的产生机制与在超导理论中的情形相似，束缚态是介质集体激发存在的原因。见 Brown and Cao(1991)。

43. 部分守恒轴矢流假说的另一个重要应用是将 π-N 散射的弹性幅和辐射幅联系起来的阿德勒规则，在这个规则中，一个额外的无质量(软)π 介子在零点能上发射。然而这个应用包含对于对称性的更多考虑，等价于南部和卢里(Nambu and Lurié，1962)的"手征性守恒"的早期思想，这一思想源于精确的 γ_5 不变性。

44. 我们区分了近似对称性和破缺对称性。前者是在经典层次上明显地被拉格朗日函数中一个小的非不变项所破坏。后者除了显然破缺的对称性外，还涉及当拉格朗日函数不变时，被非不变真空态自发破缺的对称性，以及在量子层次上被反常破坏的对称性(见 8.7 节和 10.3 节)。

45. 当电荷代数(方程(8)—(11))被扩展至时时 ETC 关系时，形式推理被认为是安全的。然而，对于时空对易式，形式规则不得不被 δ 函数的某些梯度，即所谓的施温格项所修正(见 Goto and Imamura，1955；Schwinger，1959)。至于涉及 u_0 的代数关系，见 Gell-Mann(1962a)。

46. 就威尔逊(1969)而言，这一陈述不为真。稍后将在这节和下一节给出关于其雄心勃勃的思想的更多讨论。

47. 关于这个主题的一些研究成果，几乎与约翰孙和洛的论文(1966)同时出现，这些成果是：Okubo(1966)；Bucella，Veneziano，Gatto，and Okubo(1966)；Hamprecht (1967)；Bell(1967a，b；with Berman，1967)；Polkinghorne(1967)；Brandt and Orzalesi (1967)。

48. 1966 年秋，在布朗大学召开的其时有影响力的东部理论物理学会议上，洛在"流代数的连贯性"演讲中报告了那篇论文的基本点，并发表在论文集(Feldman，1967)中。在这次演讲中，洛也强调"[ETC 关系]实际上是一个加法定则"，如果与加法定则相联系的积分是奇异的，"那么会出现反常行为"。

49. 汉普雷希特(Hamprecht，1967)和波尔金霍尼(1967)也作了分析。阿德勒和吴大峻(1969)作了最清楚的分析。见下文。

50. 用动量空间的用语，这个条件的一个等价陈述是积分变量不能正当地换元。

51. 人们不久就认可了流代数与这里和下面讨论的分歧无关。

52. 稍后的研究表明，萨瑟兰关于 $\omega \to \pi^0$ 的结论包含某些错误，并且即使在部分守恒轴矢流的框架中和流代数中也是可避免的。见 Adler(1969)。

53. 根据韦尔特曼(1991)，在韦尔特曼于 1966 年 11 月 2 日在皇家学会会议上发表演讲(Veltman，1967)之后，在韦尔特曼和贝尔之间曾有过广泛的讨论，包括电话交谈和信件交流。另见 Jackiw(1991)。

54. Bell(1967b；with Berman，1967；with Van Royen，1968；with Sutherland，1968).

55. 通过参考约翰孙和洛的论文以及其他相关论文，可以更好地理解这一陈述。

[266] 56. 出于不同的理论动机，福田和宫本(Fukuda and Miyamoto，1949)，以及施温格(1951)，都对三角形曲线图进行了相同的计算。关于这些动机的一个简要说明，见Jackiw(1972)。

57. 我很感激阿德勒教授给我提供了这封信的复印件和其他相关材料。

57a. 然而，正如巴丁所指出的(1985)："当这个调节子费米子(regulator fermion)的质量变大时，它不是退耦而是产生一个如同手征场的有效作用一样的有效'韦斯—祖米诺'作用，这已是最近通过对退耦的更一般研究而重新发现的机制。"

58. 虽然阿德勒通过他对于部分守恒轴矢流—流代数纲领的杰出贡献而赢得了他早期的声誉，但是，在他与达申(Dashen)合作并完成关于流代数的选集(1968)的编写以后，他决定做一些不同的事情。见Adler(1991)。

59. 阿德勒也注意到，由于积分线性发散，因此它的值是模糊的，并依赖于约定和计算方法。在他论文的正文中可以获得了一个不同的结果，它将满足轴向矢量沃德恒等式，但却破坏了规范不变性。见Adler(1969)脚注9。

60. 约翰孙(1963)没有借助三角形曲线图，他注意到，在非现实的然而却是精确可解的施温格模型(二维无质量量子电动力学)中，如果用一种规范不变的方式来定义流，那么拉格朗日函数的形式上的 γ_5 不变性将丧失。在这个简单模型中首次显示的定义规范不变轴矢流的不可能性，是由约翰孙提供的关于形式推理失效的另一个早期例子。

61. 在提取实验结果时，假定未实验的发散算符的矩阵元是光滑的。然而，由于方程(25)中的额外项，当矩阵元偏离 π 介子质壳时，显然不是光滑的，因而完整的发散算符不是一个光滑算符。参阅Bardeen(1969)和Jackiw(1972)。因此部分守恒轴矢流在两种用法中均以明确的方式被修正。

62. 关于与不可重正化定理相关联的历史细节，见Adler(1970)；关于其理论隐含，见Bardeen(1985)。

62a. 这在历史上是首先出现的反常取消的思想，虽然其语境不是规范理论。

63. 如果有人愿意对破坏洛伦兹不变性的路径积分的测量下定义，那么幺正性能得到拯救。

64. 与阿德勒以及贝尔和贾基夫处理的轴矢流不同，威尔逊处理的流是 $SU(3) \times SU(3)$ 流，这是在物理过程的流代数分析中出现的最为普遍的流。

65. 按照这种说法，威尔逊给瓦尔廷(Valatin，1954a，b，c，d)、齐默尔曼(1958，1967)、西岛和彦(1958)、哈格(1958)以及勃兰特(Brandt，1967)提供了参考。这种说法也为威尔逊大量吸收约翰孙的工作(1961)所证实。

66. 关于这个概念，威尔逊提到他的前辈：韦斯(1960)关于自由场论，约翰孙(1961)关于蒂林模型，以及卡斯特鲁普(Kastrup)和麦克(在Mack，1968中)关于强相互作用。此外，威尔逊也为标度假说所鼓舞，这个假说当时刚为埃萨姆和菲舍尔(1963)、菲舍尔(1964)、维多姆(1965a，b)，以及卡达诺夫(1966；及其合作者，1967)在处理统计力学中临界现象的方法方面作出了发展。参阅Wilson(1969，1983)和8.8节。

67. 威尔逊(1983)在诺贝尔奖演讲中，生动地描述了20世纪60年代中期统计物理学的进展，特别是维多姆和卡达诺夫的工作如何影响了他在理论物理学中的思考。

68. (a) 威尔逊用对蒂林模型(Wilson，1970a)和 $\lambda\phi^4$ 理论(Wilson，1970b)的分析，进一步确证了他关于场算符标度维的动力学观点。

(b) 从那时起，至少在一种工具论的意义上，甚至时空维度也已变成新的自由度。在统计物理学中，人们引入了 ε 展开技巧；在量子场论中则引入了维度规则化。这两种技巧都建立在这种时空维度的新概念基础之上。[267] 虽然现实的场论模型最终不得不是四维的(只有玩具模型可能是两维的)，然而，在统计物理学中，两维模型在真实世界中是非常适当的。

69. 值得注意的是，在威尔逊关于短程上量子场论的渐近标度不变性的概念和比约肯(1969)的概念之间有一个重要的不同。比约肯关于深度非弹性轻子—强子散射中形状因子的标度假说表明，强相互作用似乎在极端距离上停止了，而威尔逊的算符乘积展开表述，只是在把相互作用效应和重正化吸收进场和流的反常维度之后，重建了标度不变性。

但这只是表达对这个理论的标度不变进行对数校正的另一种方式。因此比约肯的思想不久就被放入非阿贝尔规范理论(量子色动力学)的框架之中，并被重新表达为渐近自由，而威尔逊的思想则在其他领域也有应用(参阅10.2节)。

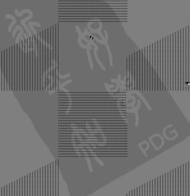
70. 温伯格(1983)也注意到，比起只是选择一个合适的能量标度，相关的问题有时更复杂，并且涉及开启集体自由度(例如强子)和关闭基本自由度(例如夸克和胶子)。

71. 威尔逊通过引入一个破坏这个理论的标度不变性的广义质量顶点，进一步推广了盖尔曼和洛的结果(1969)。卡伦(1970)和西曼齐克(1970)详细阐述了威尔逊的推广并获得了一个非齐次方程，在这个方程中，重正化群函数 ψ 是精确的，而不像在盖尔曼和洛的情形中是渐近的。当忽略非齐次项时，盖尔曼—洛方程得以重新得到。

72. 高斯不动点对应于一个自由的无质量场论；对于这个场论，场分布是高斯的。

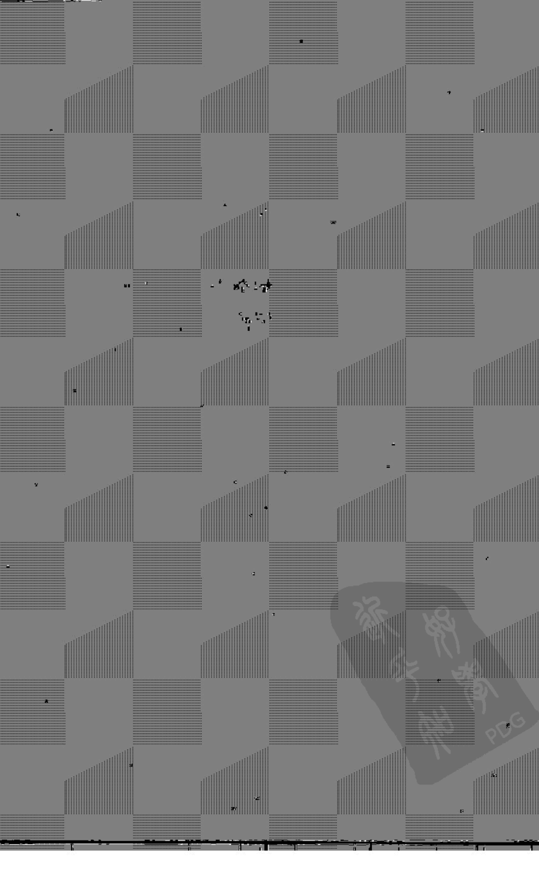
73. 正如盖尔曼和洛所指出的，量子电动力学之所以被看作微扰可重正化的，只是因为忽略了在紫外相对论性能量中微扰理论的破缺。

74. 参阅 P. D. B. Collins(1977)。



第三篇

基本相互作用的规范场纲领



现代规范场论始于杨振宁和米尔斯提出的关于强相互作用的同位 [269] 旋规范不变性理论。杨—米尔斯理论这种非阿贝尔规范理论，总体上是在量子场纲领的框架内提出的，其中相互作用由场量子传递并通过场量子之间的局域耦合实现。从物理学上讲，该理论受到强核力的电荷无关性的推动，与此同时又为这种力的短程特征所限制。从方法论上讲，它受到想拥有一个普适原理，以在众多可能性中确定唯一的耦合形式这样一种愿望的驱动。物理学家对杨—米尔斯理论产生兴趣，部分原因是他们认为它可重正化，但他们不久就放弃了这一想法，因为似乎无法拥有一个用规范不变性机制解释核力的短程特征(第 9 章)。

这个困难被如下发现所克服：首先是 20 世纪 60 年代初自发对称性破缺的发现(10.1 节)，然后是 20 世纪 70 年代初渐近自由的发现(10.2 节)。随着韦尔特曼和特霍夫特关于非阿贝尔规范理论可重正化性的证明(10.3 节)，粒子物理学共同体获得了一种似乎自洽的概念框架。

从概念上讲，这个框架在描述自然界中的各种基本相互作用，以及探索曾被认为是定域场论的新的整体性特征方面都是强有力的，而这些特征与我们对真空结构和电荷量子化的理解有直接关系(10.4 节)，从而开创了一种新的基础物理学研究纲领：规范场纲领(GFP)。

按照规范场纲领，各种基本相互作用都能用规范势(gauge potential)描述。这个纲领在所谓的标准模型中面貌焕然一新(11.1 节)。标准模型的成功鼓舞了对这个纲领的进一步扩展。人们付出 [270] 了巨大的努力，但不论是在弱电作用力和强作用力的统一上，还是在引力的规范化上，都没有成功。因此，关于规范原理的普适性问题

仍然争论不休(11.2 节)。

虽然规范场纲领的经验适当性在变迁，但是其本体论的含义还是很有趣的。关于规范场的概念是否为量子场论的几何理解提供了基础，也是可论证的。这意味着如果规范场纲领的潜能完全实现，那么规范场纲领就体现了量子场纲领和引力纲领之间的一种综合(11.3 节)。

从 20 世纪 70 年代末以来，规范场纲领没能取得进一步的进展，不论在其解释或预见基本粒子的新性质方面，还是在解决它所面临的概念困难方面。部分原因是它的停滞不前，部分原因是理论语境的改变，20 世纪场论的概念发展已经指向新的方向：有效场论(11.4 节)。

第9章
通往规范场之路

这一章主要考察规范场概念在物理学上和思想上的源头，评论 [271] 早期试图把这种富有吸引力的概念应用到各种物理过程的尝试，以及阐明这些勇敢尝试失败的原因。

9.1 规范不变性

正如我们在5.3节中所提到的，规范不变性的思想起源于1918年外尔统一引力和电磁力的尝试，这是基于四维时空中的几何方法(1918a，b)。外尔的思想如下：广义相对论除了要求坐标系只能局域地定义之外，长度标准或标度(scale)也只能局域地定义，因此需要在每一个时空点上建立一个独立的长度单位，外尔称这样一个单位标准系统为规范系统。在外尔看来，规范系统在描述物理事件时与坐标系有着同样的必要性。由于物理事件与我们对描述框架的选择无关，外尔断言规范不变性和广义协变性一样必定为任何物理理论所满足。不过外尔的标度不变性的最初想法提出不久后就被放弃了，因

为其物理含义似乎与实验相矛盾。比如，正像爱因斯坦所指出的，这个概念意味着具有确定频率的光线不可能存在。

尽管有最初的失败，但外尔的局域规范对称性思想还是保留了下来，并且随着量子力学的出现而获得了新的意义。众所周知，用哈密顿形式表述经典电磁学时，其动量 P_μ 要代之以正则动量（$P_\mu - eA_\mu/c$）。这种替代需要把电磁相互作用构造成经典运动方程，[272] 也就是要确定电磁相互作用的形式。正则动量的形式是受哈密顿原理启发而来。这里的基本思想是想从单一的物理学原理同时得到麦克斯韦方程组和带电粒子的运动方程。量子力学中的一个关键概念是经典哈密顿量中的动量 P_μ 被算符（$-\mathrm{i}h\partial/\partial x_\mu$）所代替。这样一来，正如福克（Vladimir Fock，1926）所注意到的，量子电动力学可以建立在正则动量算符 $-\mathrm{i}h(\partial/\partial x_\mu - \mathrm{i}eA_\mu/hc)$ 的基础上。不久伦敦（Fritz London，1927）指出了福克的工作和外尔的早期工作之间的相似性。外尔的规范不变性思想，如果其标度因子 ϕ_μ 被（$-\mathrm{i}eA_\mu/hc$）替代，那么就将是正确的。因此，我们考虑的是用相变 $[1-(\mathrm{i}e/hc)A_\mu\mathrm{d}x_\mu] \approx \exp[-(\mathrm{i}e/hc)A_\mu\mathrm{d}x_\mu]$ 取代标度变化（$1+\phi_\mu\mathrm{d}x_\mu$），并且相变可以看作虚的标度变化。在外尔 1929 年的论文中，他把所有这些考虑结合起来，并明确讨论了电磁势的变换 $A_\mu \to A'_\mu = A_\mu + \partial_\mu\Lambda$（第二类规范变换），以及带电粒子波函数的相关的相变换 $\Psi \to \Psi' = \Psi\exp(\mathrm{i}e\Lambda/hc)$（第一类规范变换）。

这里，主要的线索由认识到波函数的相可能是一个新的局域变量而提供。先前对外尔的最初思想的反驳不再适用，因为波函数的相不直接包含在对诸如一个矢量的长度这类时空量的测量中。因此，在电磁场不存在的情况下，相的变化量可取任意常量值，因为这不会影响任何可观测量。当有电磁场存在时，在每一时空点的相的不同

选取也容易通过把势 A_μ 解释成联络(connection)而协调起来，这种联络把不同点处的相联结在一起。相函数的特定选择不会影响运动方程，因为相中的变化和势中的变化恰好彼此抵消。按照这种方式，从前起因于势的明显“任意性”，现在就可以理解为选择不影响方程的波函数相值的自由性。这正是规范不变性所确切表示的意思。

在把外尔的成果应用到量子场论时，泡利曾在其颇有影响的评论性论文(1941)中指出，整体相变(Λ 不依赖于其时空位置的第一类规范变换)下的理论不变性要求电荷守恒，而局域规范不变性(局域的含义是指 Λ 依赖于其时空位置)与电磁相互作用有关，并且通过代换 $[\partial/\partial x_\mu \rightarrow (\partial/\partial x_\mu - \mathrm{i}eA_\mu/hc)]$ 决定了它的形式，即所谓的最小电磁耦合。

有三点现在是清楚的。第一，没有量子力学就不可能得到规范不变性的正确解释。第二，规范理论与广义相对论在理论结构上有 [273] 很大程度的相似性，即一边是电磁场、局域相函数、电磁势和规范不变性，另一边是引力场、局域切矢量、仿射联络和广义协变性。

第三，正如迈克·弗里德曼(1983)已经指出的，不变性的概念不同于协变性的概念。如果某些几何对象在群的作用下是不变的，那么这个理论在变换群下就是不变的；如果在变换群的作用下理论中方程的形式是不变的，那么这个理论就是协变的。考虑到这个细微的区别，规范不变性的概念在许多情况下实际上都是指规范协变性，并且规范理论中涉及的对称群实际上是协变群，而不是不变群。[1]顺便指出，这个区分适用于本书的其余所有地方。通常不变性和对称性是在协变性的意义上使用。比如，在相关的可观测自由度的不同表象之间的变换下，如果在变换群的作用下理论中的方程或定律的形式

是不变的，那么理论就被说成是对称的或不变的。 但当我们说只有规范不变的那些量在物理上是可观测的，我们是在“不变性”本来的意义上使用规范不变性的概念。

9.2 确定相互作用形式的规范原理

规范理论的现代思想始于杨振宁和米尔斯 1954 年发表的两篇论文(1954a，b)中的思索。 他们思索的核心或者说任何规范理论的核心，都是用以决定理论的动力学的规范对称群。 为了帮助理解，我们先给出一些定义。 如果一种对称性的表示在每一时空点都是相同的，那么这种对称性是整体的；如果一种对称性的表示在不同时空点彼此不同，那么这种对称性是局域的。 因此，洛伦兹对称性是整体对称性，而广义坐标变换对称性则是局域对称性。 再者，如果相关可观测自由度在本质上是时空的，则这种对称性为外部的，否则称为内部的。 因此，洛伦兹对称性是外部对称性，而量子电动力学中的相位对称性则是一种内部对称性。

杨—米尔斯理论完全兴起于量子场纲领的框架之内，并且受到以下两种思考的激发。 第一，在第二次世界大战后人们发现了越来越多的新粒子，并且提出了这些基本粒子之间的各种可能耦合。 因此[274] 杨振宁和米尔斯感到有必要找到某一原理，从而可以从所考虑的众多可能性中选择唯一的一种形式。 杨振宁和米尔斯提出的这个原理基于规范不变性的概念，因而被称为规范原理(gauge principle)。 第二，在选取恰当的规范对称性时，杨振宁和米尔斯受好奇心的驱使，要寻找出若假定有同位旋守恒定律将会有什么样的结果，这种假定是把强相互作用与电荷相互作用进行类比的产物。[2]

在数学上，同位旋守恒意味着局域相位对称性向一般局域化的内

在对称性的推广。在物理上，它是对核力的电荷无关性这一经验发现的重新表述。仿效海森伯，该理论假定质子和中子仅仅是同一种粒子的抽象同位旋空间中的两个状态，于是在量子电动力学中电荷守恒就与相位不变性有关。因此，通过类比，人们就可假定强相互作用在同位旋旋转下有不变性。同位旋不变性要求同位旋的方向不是物理意义上的方向。中子和质子之间的不同便是一种完全任意的选择过程。然而，这种任意性具有整体特性，也就是说，一旦在一个时空点上作出了某种选择，那么在所有其他点上的选择就都固定了下来。但是正如杨振宁和米尔斯认识到的，这与量子场论中每个场量只能局域定义的要求不一致。

为了得到一个局域同位旋不变的理论，杨振宁和米尔斯受与量子电动力学类比的引导。在量子电动力学中，局域相位不变性是通过引入电磁势 A_μ 得以保持，A_μ 抵消带电粒子的波函数的相位变化。杨振宁和米尔斯以一种完全类似的方法，引入了规范势 B_μ，在同位旋旋转变换 $B_\mu \rightarrow B'_\mu = S^{-1}B_\mu S + (\mathrm{i}/\epsilon)S^{-1}\partial S/\partial x_\mu$ 下，用 B_μ 来抵消核子波函数的同位旋旋转 S 对时空的依赖。这里 $S = \exp(-\mathrm{i}\theta(x)\boldsymbol{L})$ 是局域同位旋群的一个元素，这种同位旋群是非阿贝尔李群：$\boldsymbol{L}$ 有三个分量 L_i，L_i 是群的生成元。因此，在杨—米尔斯理论中，强相互作用中与电荷类似的是同位旋，而与光子相类似的是三维矢量规范玻色子 B^i_μ。杨振宁和米尔斯所获得的这些结果意义重大。

第一，容易证明势 B_μ 一定包含 L_i 的线性组合：$B_\mu = 2\boldsymbol{b}_\mu \cdot \boldsymbol{L}$。这个关系明确显示了杨—米尔斯势既作为时空中的四维矢势又作为 SU(2) 同位旋矢量算符的双重作用。也就是说，杨—米尔斯势是非阿贝尔势。当考虑比如来自 SU(3) 色同位旋的内禀量子数时，推广至其他非阿贝尔势在概念上就是很直接的。

[275] 第二，势 B_μ 的十二个独立分量 b_μ^i 所满足的场方程和它们与具有同位旋荷的任何场的相互作用，在本质上都被同位旋不变性的要求所确定，同样地，电磁势 A_μ 和 A_μ 与电场的相互作用，在本质上也是由相位不变性的要求所决定。因此，我们就有了决定基本相互作用形式的强有力原理，即规范原理。这个为量子场论的旧表述所缺少的原理是规范场纲领的奠基石。

9.3 早期努力

杨振宁和米尔斯提出他们关于同位旋不变性场论的思索，并试图用它解释强相互作用。这种规范不变性理论明显的诱人之处在于它可能是可重正化的。但是就强相互作用而言，这个理论明显是站不住脚的。正如汤川秀澍所证明的，强相互作用的有限量程要求它由非零质量的量子来传递。但是有质量量子也会破坏规范不变性和可重正化性。另一方面，正如泡利指出的，为了保留规范对称性，其规范玻色子必须是无质量的[3]，但是这样一来，这些规范玻色子对短程相互作用就没有什么用处。

1956 年，内山菱友(Ryoyu Utiyama)通过在更宽泛的对称群下假设系统的局域规范不变性，以确定的方式引入一种更一般的法则，提出了一种跟源场有确定类型的相互作用的新场。正当杨振宁和米尔斯还将注意力局限于内部对称性时，内山菱友的处理方法已经延伸到外部对称性。这样，按照内山菱友的观点，引力场和引力相互作用也可纳入到规范理论的一般框架。内山菱友的想法引起很多争议[4]，但在量子引力上也得到一些支持(见 11.2 节)。

施温格(1957)作出了更具成效的努力，他基于 SU(2)群提出了一个杨—米尔斯类型的规范理论，在其中电磁相互作用和弱相互作用能

按一种统一的方式加以解释：中性矢量规范玻色子(光子)被假定为与电磁流耦合，而带电矢量规范玻色子($W^\pm$)被假定为与守恒弱流耦合，形成SU(2)对称群的一个三重态。三重态中无质量的光子和有质量的$W^\pm$玻色子之间存在的巨大质量分裂，则通过引入辅助标量场加以解释。

施温格的想法被布卢德曼(1958)、萨拉姆和沃德(1959)以及格拉肖(1961)等人所采纳，后来也由萨拉姆和沃德(1964)、温伯格(1967b)和萨拉姆(1968)加以详细阐述。 [276]

布卢德曼的工作重点在以下三方面。第一，他把同位旋对称性扩展到手征SU(2)对称性，以使弱相互作用中的宇称不守恒得到解释。作为这项工作的额外收获，他成为第一个提出"电荷固定的费米相互作用"(后来被重新命名为中性流)的物理学家，这是除传统的电荷交换的相互作用之外的相互作用。第二，他接受施温格的想法，认为像强相互作用中的强子一样，弱衰变中的轻子携带弱同位旋；并且，他为了实现这个想法，赋予已知轻子手征SU(2)群一种双重态表示。这种赋予方法在后来的研究者中沿袭了下来，并成为标准模型的一部分。第三，他提出带电的或中性的矢量玻色子可以用有质量杨—米尔斯场描述。他的这些意见开启了研究的新领域，即有质量杨—米尔斯理论。布卢德曼也观测到带电矢量玻色子有奇异的、不可重正化的电磁相互作用。这一观测结果激发了对具有带电规范玻色子的规范理论的可重正化性的深入研究。[5]

格拉肖(1961)沿着施温格的足迹继续寻求弱相互作用和电磁相互作用的统一。不过他认为在SU(2)群结构中，不可能按一种统一的方法描述宇称守恒的电磁流和宇称破坏的弱流。因此他扩展了布卢德曼的工作，而在布卢德曼的工作中，SU(2)对称性只对弱相互作用

起作用；但除了 SU(2)之外，布卢德曼还通过引入一个阿贝尔群 U(1)而把电磁作用包括进去。因此基本群结构应该是 SU(2) × U(1)，并且轻子的相互作用的最终拉格朗日函数有四个矢量玻色子：两个 $W^{\pm}$ 玻色子和带电弱流耦合，一个 W^{0} 玻色子和中性弱流耦合，一个光子和电磁流耦合。在四维矢量玻色子中，格拉肖也引进了一个光子—W^{0} 玻色子混合角，后来被称为温伯格角。从美学角度看，格拉肖的模型很精彩。从现象学的角度看以及从后来的发展判断，它为统一电磁相互作用和弱相互作用给出了最好的模型，并且得到实验的有力支持。不过，它却没有论及杨—米尔斯类型理论的一个基本理论困难，即有质量规范玻色子为什么没有破坏规范不变性。

自发对称性破缺和施温格—安德森—希格斯机制(见 10.1 节)的发现才解决了这个困难，这个机制被纳入到温伯格(1967)和萨拉姆

[277] (1968)提出的模型中。然而对所有这些模型来说，非阿贝尔规范理论仍有一个基本困难，即仍没有谈到它们的可重正化性问题。温伯格曾推断他的模型因为规范不变性可能是可重正化的。然而因为在 10.3 节中才清楚的原因，这种推断在技术上是站不住脚的并且在概念上是空洞的。

具有杨—米尔斯类型理论结构的一个有影响的强相互作用理论，是樱井纯 1960 年提出的。樱井是第一个从具体物理模型而不是用纯粹推测去探索强相互作用的规范原理重要地位的理论物理学家。他提出强相互作用要通过五个有质量矢量规范玻色子作为介质，这跟规范群 SU(2) × U(1) × U(1)有关，并且因此与三个守恒的矢量流耦合，它们是：同位旋流、超荷流和重子流。

由于熟知质量困难，樱井寻找动力学机制来产生质量。一种可能性是从规范玻色子的自相互作用中产生质量。他写道："希望能

够从严格遵守我们的规范原理的基本拉格朗日函数产生这种似乎违背了我们的规范原理的有效质量项，并非完全是荒唐的。"（1960）这句话明确预料到对称性的自发破缺思想和所谓的希格斯机制，这在后来使规范理论得以给出了关于弱电相互作用的一个前后一致的和成功的解释。[6]

樱井关于自发对称性破缺和希格斯机制的预期是勇敢的一步。他在其论文的一个脚注中写道："这一点[关于从拉格朗日函数中产生质量]已经被贝伦兹(R. E. Behrends)博士、奥本海默教授和派斯教授非常严厉地批评过。"（同上）另外一个注释更有趣：

> 对我们理论的几种批评认为，由于我们不大可能成功解决质量问题，我们最好还是把具有质量的 B 场[的方程]取作理论的出发点，而把和规范原理的可能关联置之脑后。这种手法几乎可以适应所有的实际目标。然而，作者相信任何理论上的努力都是为了在先验理论基础上去证明基本耦合。
>
> （同上）

在这些批评者中有盖尔曼。最初盖尔曼他自己被杨—米尔斯理论所吸引，并试图找到一个"软质量"机制让假想的无质量理论的"可重正化性"保留下来。但他失败了。正如他后来回忆的(1987)，他在20世纪60年代就对杨—米尔斯理论失去兴趣，并且甚至劝说包括樱井在内的其他规范理论学家放弃强相互作用中的规范原理，转而采用"矢量介子支配"模型——一种更"实际"的可选方案。然而，盖尔曼早期对杨—米尔斯理论的迷恋不是没有结果的。理论上，他的流代数只有在杨—米尔斯理论的框架下才是可理解的， [278]

其中矢量(和/或轴矢)玻色子是与费米子的矢量(和/或轴矢)流耦合在一起的。其对易规则是规范不变理论的对称性的结果。因不能解决规范玻色子质量问题而只得去除玻色子，所剩的就只是满足流对易规则的流了。

流代数和杨—米尔斯理论之间的逻辑关联，在20世纪60年代就被一些物理学家敏锐地觉察到，并得到富有成果的利用；当时流代数取得许多有意义的成就，其中最引人注目的是阿德勒—韦斯伯格关系(Adler-Weisberger relation)。比如，韦尔特曼(1966)用他所谓的发散方程组成功地导出阿德勒—韦斯伯格关系和流代数的其他结果以及部分守恒轴矢流纲领。在发散方程组中，流代数的方程组和部分守恒轴矢流纲领通过用 $(\partial_\mu - W_\mu \times)$ 替代 ∂_μ 而扩展到包括高阶弱效应和电磁效应，其中 W_μ 代表矢量玻色子。作为对韦尔特曼工作的回应，贝尔(1967b)从规范变换的思考出发，提出离散方程的一个形式推导，因此有一点是清楚的，即流代数和部分守恒轴矢流纲领的成功一定是规范不变性的结果，至少对于像韦尔特曼那样有场论取向的物理学家来说是这样的。察觉到流代数和杨—米尔斯理论之间的这种逻辑关联，韦尔特曼(1968b)成功地取得了以阿德勒—韦斯伯格关系作为非阿贝尔规范理论可重正化性的经验证据；其中有一个隐蔽的假设，即只有可重正化理论才可以成功地描述自然界中所发生的事情。正如我们将在10.3节中所见的，韦尔特曼对杨—米尔斯理论的可重正化性的信心产生了重要后果。

注释

1. 弗里德曼也把理论的对称群定义为标准形式的协变群，后者只是(理论的)动力学客体的微分方程系统。由于在规范理论中所有客体都是动力学的，所以对称群和协变群

是简并的。

2. 参阅 C. N. Yang(1977，1983)。

3. 规范玻色子无质量的理由如下：为了从欧拉—拉格朗日方程得到矢量规范场的正确的运动方程，势的质量一定要通过形式为 $m^2 A_\mu A^\mu$ 的项引入到拉格朗日函数中。这一项显然不是规范不变的，因为由势的变换出现的附加项没有被费米子的波函数的变换所抵消。因此，在杨—米尔斯理论中或者更普遍地在任何规范理论中，都要求规范场无质量。这是局域规范不变性无法避免的结果。 [279]

4. 例如，见基布尔(T. W. B. Kibble，1961)和 Y. M. Cho(1975，1976)之间的争论。关于外部规范对称性的更多内容，另见 Hehl et al. (1976)，Carmeli(1976，1977)和 Nissani(1984)。

5. 例如，见 Glashow(1959)，Salam and Ward(1959)，Salam(1960)，Kamefuchi(1960)，Komar and Salam(1960)，Umezawa and Kamefuchi(1961)以及其他许多论文。

6. 樱井的论文(1960)是在有关对称性的自发破缺的其他论文之前向出版机构提交的，比 Nambu and Jona-Lasinio(1967a，b)和 Goldstone(1961)更早。在 10.1 节中有关于自发对称性破缺的更多讨论。

第 10 章
规范场论概念基础的形成

[280] 量子场的规范不变性体系，作为描述各种相互作用的自洽框架，必须要找到短程相互作用的某些机制(10.1 节和 10.2 节)及其可重正化的证明(10.3 节)。另外，非阿贝尔规范理论呈现出一些新特征，这些新特征提出了关于真空态结构的明确解释，而且提出了诸如电荷之类的物理参量的量子化条件。因此一个从没有在(阿贝尔规范不变的)量子电动力学或其他非规范不变的定域场论的传统基础研究中出现过的新问题，颇为紧迫地暴露了出来，并引起强烈关注，成为一个为数不少的数理学家群体近年来钟爱的研究题目。这就是非阿贝尔规范场论的整体特征问题(10.4 节)。这一章将回顾规范理论(它们既是理论框架又是研究纲领)的这些概念基础的形成。这一章还提到了一些有待未来研究者解决的问题。

10.1 短程相互作用的机制(I)：自发对称性破缺

最初的杨—米尔斯理论未能对强核相互作用的已有理论有所改

进，因为杨—米尔斯理论不能在显然不违背规范对称性的前提下，再现核力的可观测的短程行为。因而，在规范理论的进一步发展中，需要克服的一个主要困难是要提出一个自洽的方案，这个自洽的方案在（某种意义上）保持规范不变性的同时还具有质量不为零的规范量子。解决这个问题的一条途径就是所谓的自发对称性破缺（spontaneous symmetry breaking，SSB）。[1]

[281] 什么是自发对称性破缺？假定一个拉格朗日函数在对称群 G 作用下是不变的，并且存在一个构成 G 多重态的最低能量状态（基态）的简并集。如果简并集中的某个基态被选作系统的"当前"基态，那么系统的对称性就被说成自发破缺的。在这种情况下，就所涉及的基本定律而言，对称性仍然存在。但是在系统的实际状态中，对称性并没有显现出来。也就是说，对称的物理理论在物理表现上是很不对称的。我们能在铁磁体现象中找到自发对称性破缺的一个简单例子：虽然铁磁体系的哈密顿函数具有旋转不变性，但是系统的基态具有沿某个任意方向排成一线的电子自旋，因此表现出旋转对称性的破缺，并且任何建立在基态上的高级态也有它的非对称性。

从 20 世纪 50 年代末以来，人们对某些量子场系统可以显示自发对称性破缺这一思想作了明确研究；最初是南部本人（1959），然后是他和同事约娜-拉西尼奥（G. Jona-Lasinio，1961a，b），随后还有戈德斯通（Jeffrey Goldstone，1961）和其他一些人。因为这个原因，许多物理学家倾向于认为自发对称性破缺在 1960 年左右上述作者的文献中才初露端倪。其实，自发对称性破缺的思想比粒子物理学还要早得多，在被重新发现和在 1960 年左右融入规范理论框架之前，它已经有很长一段历史了；具体说来，它产生于凝聚态物理学和粒子物理学的迷人发展之间的一种积极的相互影响。

自发对称性破缺的早期思想

现代对物理学中对称性的重要性的认识始于赖伊[和恩格尔(Engel), 1893]以及皮埃尔·居里(Pierre Curie, 1894)。赖伊主要关心自然律(即方程)的对称性，而居里感兴趣的是物理状态(即方程式的解)的对称性。[2]居里没有现代的自发对称性破缺思想，但他认识到会发生某些现象。他说“某些对称性要素的丧失是必然的”，甚至宣称“非对称性创造现象”(同上)。自发对称性破缺的经典例子很多；其中有些甚至早在居里1894年的论文之前就为人所知。这些自发对称性破缺例子的一个共同特点是，它们产生于非线性系统，这种系统对参量的某个关键值有分岔点，比如流体动力学中的雷诺数，或旋转的自重流体问题中的角动量的平方。

在统计热力学的语境中，庞加莱(1902)在讨论遵守牛顿时间可逆 [282] 定律的微观现象如何能导致明显不可逆的宏观现象时，在更深层次上指出自发对称性破缺的另一面：系统必定由很多要素组成。因此，虽然系统“循环”足够长的时间后可随意接近其初始状态，但是正如遍历定理(ergodic theorem)所要求的，在实际的大物理系统中“庞加莱循环”时间是不可实现地长。

海森伯关于铁磁性的论文(1928)提出了量子理论中自发对称性破缺的现代经典阐释。他为一对电子自旋的相互作用能假定了一个明显的对称形式，即

$$H_{ij} = J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j$$

(其中$\boldsymbol{S}_i$和$\boldsymbol{S}_j$是铁磁体的最近邻格点上原子的自旋算符)。然而，在SU(2)自旋旋转群之下的不变性被磁场干扰所破坏，导致所有自旋方向都指向同样方向的基态。这个状态是用宏观参量$\langle \boldsymbol{S} \rangle$(磁化强度)描述的。虽然存在无穷多种可选方向，在该方向上所有自旋都排

成一线，并且它们的能量都相同，但是由于所考察的系统太大，以至于既没有量子涨落也没有热涨落可引起在任何可想象的有限时间内向不同基态的跃迁。因此与铁磁体相关的自发对称性破缺现象能从“大数论证”（large-number argument）的角度来理解，并为宏观参量即磁化强度〈**S**〉的非零值所刻画。[4]

1937年，朗道结合其相变理论（1937）推广了非零的宏观对称性破缺参量的理论，尤其是后来又结合他和金兹堡（V. L. Ginzburg）得到的超导理论再作推广（1950）。朗道的工作使自发对称性破缺重见光明。首先，在关于连续相变的讨论中，朗道证明，无论何时只要不同的相具有不同的对称性，自发对称性破缺都会发生，因此建立了自发对称性破缺的普适性；他还揭示了其物理原因：系统在寻求能量最佳状态。其次，他引入了特征序参量概念，这种参量在对称相中变为零而在非对称相中不为零。

在超导的情况下，朗道引入超导电子的“有效波函数”Ψ作为特征函数，并写出了唯象论方程来描述它，得到了重要的结论：如果Ψ不为零，那么自发对称性破缺一定发生，并且用Ψ描述的超导态是非对称的。

在1958年11月21日，朗道出于对海森伯的非线性统一场论的支持，写道：

> 方程的解将有比方程本身程度低一些的对称性。电磁相互作用或许还有弱相互作用都是从非对称性中得出的……这更像一个纲领而非理论。这个纲领虽然宏大，但是仍必须贯彻。我相信这将是理论物理学的主要任务。[5]

[283]

海森伯的真空新概念

在20世纪50年代末，海森伯对自发对称性破缺的重新发现作出了至关重要的贡献。[6]海森伯的非线性统一场论纲领的关键问题是，如何从基本场方程导出不同相互作用领域中的各种现象，这些相互作用拥有不同的动力学对称性，而这些方程拥有比现象自身更高的对称性。为了处理这个问题，海森伯把注意力转向真空概念，真空的性质奠定了任何场论的概念结构的基础。

1958年，高能核物理学的罗切斯特会议在日内瓦举行，海森伯借用简并真空的概念来解释内禀量子数，比如同位旋和奇异数，这些量子数提供了基本粒子相互作用的选择定则(1958)。[7]

在一篇1959年提交的很有影响的论文中[8]，海森伯和他的合作者用他在量子场论中的简并真空概念，解释了电磁相互作用和弱相互作用引起的同位旋对称性的破缺。他们说"没有办法能先验地确定[这个理论]必定给出一个拥有初始方程的所有对称属性的真空态"(同上)。而且，"当明显不可能构造一个完全对称的'真空'态时"，应该"实际上与其看作'真空'倒不如看作'世界态'，这种'世界态'形成基本粒子存在的亚态，而这个状态一定是简并的"，并且可能"在目前的理论中有一个实际上是无穷大的同位旋"。最后，他们把这个"世界态"作为"对称性破缺的基础，例如，根据量子电动力学，在同位旋群中的破缺正是在实验上所期望的[因为相关的激发态能够从'世界态'中'借出'同位旋]"(同上)。

海森伯的简并真空立刻在国际会议上得到广泛讨论。[9]它被频繁引用，极大地影响了场论学家，并有助于扫清自发对称性破缺从流体动力学和凝聚态理论延伸到量子场论的道路。不过，海森伯仍没有得到对自发对称性破缺的起源、机制和物理结果的满意理解，他也没

有给出令人信服的数学形式。[10]

[284]

来自超导性的鼓舞

20 世纪 50 年代后期超导理论的新进展，鼓舞人们寻求理解自发对称性破缺的更有效途径。为了解释超导体中所测到的比热容和热导率在低温时变得很小的事实，巴丁(1955)提出一个在能谱中仅高于最低能值的能隙模型。不久，他与合作者库珀(Leon N. Cooper)和施里弗(J. Robert Schrieffer)把这个模型发展成超导性的微观理论，即 BCS 理论(1957)。在此理论中，能隙产生于明确的微观机制，即库珀对：结合在一起的电子库柏对，具有方向相反的动量和自旋，以及接近费米面的能量，这些都是由电子之间以声子作为介质的相互吸引作用引起的；并且"能隙"就是打破这种结合所需的具体能量大小。

规范不变性和 BCS 理论

虽然 BCS 理论能给出超导体的热力学属性和电磁属性的定量解释，包括迈斯纳效应(Meissner effect，超导体内部磁流的完全溢出)的解释，但是这只能在特定的规范中，即朗道规范(具有 $\mathrm{div}\,A = 0$)中才能做到。要是 BCS 理论是规范不变的，在特定规范中进行这种运算应该没有问题，但是情况并非如此。因此，从 BCS 理论推导出迈斯纳效应，以及 BCS 理论本身，成为 20 世纪 50 年代后期激烈争论的主题。[11]

巴丁、安德森(P. W. Anderson)、派尼斯(Pavid Pines)、施里弗和其他人试图恢复 BCS 理论的规范不变性，与此同时保留能隙和迈斯纳效应。[12]巴丁在争论中提出一个重要看法，即能隙模型和规范不变理论之间的差异仅仅在于静电势的规范不变理论中所蕴含的东西。这种蕴含的东西可能产生纵向集体模，但却跟与横向矢势相联系的磁

场相互作用没有什么关系。

安德森和其他人精炼了巴丁的论点。安德森证明了纵向型集体激发恢复了规范不变性。如果把长程库仑力包括在内，则纵向激发类似于高频等离子体，并且没有填满零点能附近的能隙。安德森的主张可以重新表述为：(i)规范不变性要求无质量集体模的存在；以及(ii)长程力把这些无质量模再结合成有质量模。借用到量子场论中，第(i)点对应于南部—戈德斯通模，而第(ii)点对应于希格斯机制。不过，使这种主张首次成为超导性和量子场论之间的一座桥梁以及构建这座桥梁，很大程度上是由南部完成的。

[285]

通向 BCS 理论规范不变性的南部进路

20 世纪 50 年代后期，南部采用了博戈留波夫(Bogoliubov)对 BCS 理论的重新表述，并把它看作广义哈特里—福克近似进行处理[13]，从而澄清了规范不变性、能隙和集体激发之间的逻辑关系。更重要的是，南部提供了一个探索超导性和量子场论之间富有成效的类比的恰当起点。

博戈留波夫对超导性的处理，把超导体中的元激发描述为电子和空穴(准粒子)的相干叠加，这些电子和空穴遵守博戈留波夫—瓦拉京方程(Bogoliubov-Valatin equations)：

$$E\Psi_{p+} = \varepsilon_p \Psi_{p+} + \phi \Psi^{+}_{(-p)-}, \quad E\Psi^{+}_{p-} = -\varepsilon_p \Psi^{+}_{(-p)-} + \phi \Psi_{p+} \quad (1)$$

其中 $E = (\varepsilon_p^2 + \phi^2)^{1/2}$。$\Psi_{p+}$ 和 Ψ_{p-} 是动量为 p、自旋为 + 或为 − 的电子的波函数，因此 $\Psi^{+}_{(-p)-}$ 实际上表示动量为 p、自旋为 + 的空穴；ε_p 是在费米面上测得的动能；ϕ 是源自以声子为介质的电子之间吸引力的能隙。这个理论不可能是规范不变的，因为准粒子不是电荷的本征态。不过，南部的工作表明，规范不变性的缺失并不是

BCS—博戈留波夫理论的缺陷，但是由于能隙是规范相关的，因此规范不变性的缺失深深地植根于超导体的物理实在中。[14]

南部系统地把量子动力学方法应用到BCS—博戈留波夫理论。他写下自能部分的方程，这对应于哈特里—福克近似，并且还得到一个非微扰解，其中规范相关的 ϕ 场最后证明就是能隙。他还写下顶点部分的方程，并发现如果把“辐射修正”包括在内，就与自能部分相关，因此建立了“沃德恒等式”和理论的规范不变性。他得出一个重要发现：顶点方程的精确解之一是具有零能量和动量的粒子对的束缚态，这个解即准粒子对的集体激发，正是它导致“沃德恒等式”。

在南部关于超导性的自洽的物理绘景中，准粒子总是伴随着周围介质的极化(辐射修正)，并且如果把两者结合起来就会导致电荷守恒 [286] 和规范不变性。极化场的量子表明，量子自身是介质的集体激发，而介质由运动的准粒子对构成。因此这些集体激发(束缚态)的存在表现为规范不变性的逻辑结果，是与规范相关的能隙的存在相匹配的。

从超导性到量子场论

南部在超导性方面的工作使他想到，有可能把非不变解的思想用在粒子物理学上(尤其在真空态中)。南部在1959年的基辅会议上第一次表述了这个重要观点。

> 在场论的 γ_5 不变性问题和博戈留波夫表述的超导规范不变性问题之间[存在]一种类比。观测到的粒子的质量在这种类比中对应于超导态中的能隙。超导性的这种巴丁—博戈留波夫描述(Bardeen-Bogoliubov description)，由于能隙的原因而不是规范

不变的。但已有人成功地按规范不变的方法解释了巴丁—博戈留波夫理论。按照这种方法，人们也可以处理 γ_5 不变性。因此，例如，人们容易得出结论，认为在可能和 π 介子态是相同的赝标量态下，存在束缚核子—反核子对。

(1960)

一年后，南部在 1960 年的美国中西部物理学会议上，用下面的对应表总结了这个类比：

超导性	基本粒子
自由电子	裸费米子(零质量或小质量)
声子相互作用	一些未知的相互作用
能隙	观测到的核子质量
集体激发	介子(束缚核子对)
电荷	手征性
规范不变性	γ_5 不变性(精确的或近似的)

在数学层面上，这种类比是完备的。在博戈留波夫—瓦拉京方程组(方程组(1))和 γ_5 不变的狄拉克理论中的方程组之间有严格的相似性。γ_5 不变的狄拉克理论中的方程组为：

$$E\Psi_1 = \sigma \cdot p\Psi_1 + m\Psi_2, \quad E\Psi_2 = -\sigma \cdot p\Psi_2 + m\Psi_1 \qquad (2)$$

[287] 其中 $E = \pm(p^2 + m^2)^{1/2}$，并且 Ψ_1 和 Ψ_2 是手征算符 γ_5 的两个本征态。通过利用广义哈特里—福克近似，核子质量 m 能够作为一个原本无质量核子的自质量来获得，而正是这个质量损坏了 γ_5 不变性。[15] 为了恢复这个不变性，就要求无质量的集体态，这种态可解释为核子对的赝标量束缚态，即基本场的自相互作用所产生的一种效应。[16]

注意一下狄拉克和海森伯对南部寻求这种类比的冲击力，是有趣的。首先，南部认真对待狄拉克的空穴思想，并且不把真空看作一

种虚空，而是看作一种塞满许多虚自由度的充盈空间。[17]这种充盈真空观使南部有可能接受海森伯关于真空简并性的概念，这个概念正是自发对称性破缺的核心。[18]其次，南部试图构建一个复合粒子模型，并且选择了海森伯的非线性模型，“因为对称性破缺的数学面貌在此已充分显露”，虽然他从不喜欢这个理论甚至也没有认真对待过它。[19]由于注意到超导性的博戈留波夫表述和海森伯的非线性场论在数学上的相似之处，南部引入了π介子—核子系统的特殊模型，把从超导理论中得到的有关自发对称性破缺几乎所有的结果，引进到量子场论中。这些结果包括非对称(超导类型)的解、真空态的简并和恢复对称性的集体激发。

这些平行概念的一个例外是等离体子(plasmon)，等离体子在超导理论中源自集体激发和长程力的综合效应。在南部的场论模型中，等离体子没有类比物，一方面是因为π介子衰变施加的限制[20]，另一方面是没有类似的长程力存在。然而，南部预料到他的理论可能对解决杨—米尔斯理论的困境有帮助，因为他写道：

> 在这种类型的理论中，有另一个令人着迷的问题，即能否在规范不变理论中使玻色子产生有限的有效质量。因为我们已经在费米子上做到了，如果能把一些额外自由度加到场上，这应该不是不可能的。倘若这个答案最终是肯定的，那么关于矢量玻色子的杨—米尔斯—李—樱井理论可能的确非常有意义。
>
> (1960c)

自发对称性破缺进入规范理论框架的最后整合

南部关于自发对称性破缺的工作，主要依赖于海森伯非线性场论

和博戈留波夫的超导理论方案之间的类比。但是这个类比并不完美，因为超导性中的能隙与费米球相关，所以是一个非相对论性概念。[288] 为了完善这个类比，在海森伯相对论性不变的非线性理论中要求有个动量截止。在海森伯理论中需要一个截止的另一个原因是该理论不可重正化。因此，下一步应当是寻找一个可重正化模型，其中自发对称性破缺可以在没有任何截止的情况下进行讨论，这是超越类比的一步。

超越类比

戈德斯通在读了南部的论文预印本后，很快完成了这一步（1961）。戈德斯通的工作推广了南部的结论。戈德斯通放弃非线性旋量理论，考虑了只有一个复玻色子场中的简单的可重正化量子场论，其拉格朗日密度为

$$L = (\partial \phi^* / \partial x_\mu)(\partial \phi / \partial x_\mu) - V[\phi], \quad (3)$$

其中 $V[\phi] = \mu^2 \phi^* \phi - (\lambda/6)(\phi^* \phi)^2$。在 $\mu^2 < 0$ 且 $\lambda > 0$ 时，势能项 $V[\phi]$ 在由 $|\phi|^2 = -3\mu^2/\lambda$ 决定的复平面 ϕ 中有一条底线。这些最小值是真空态中玻色子场可能的非零期望，表明真空态是无限简并的。戈德斯通称这些真空解是异常解，或叫超导解。显然，含有 λ 的四次项对得到玻色子场的非零真空期望来说是至关重要的，按照我们的理解，正是非线性相互作用为稳定的非对称解的存在提供了动力学机制。

在 $\phi \to \exp(\mathrm{i}\alpha)\phi$ 变换下，拉格朗日方程有明显的不变性，并且方程的解能通过对相 α 的赋值而得以确定。戈德斯通通过禁止建立在不同真空上的态的叠加，而在相上设置了超选择定则。因此人们必须选择某一特殊的相，从而破坏了对称性。如果我们用 $\phi = \phi' +$

χ 定义一个新场 ϕ'，χ 是实数并且 $\chi^2 = |\phi|^2 = -3\mu^2/\lambda$，$\phi' = 2^{-1/2}(\phi_1' + i\phi_2')$，则拉格朗日函数变成：

$$L = 1/2[(\partial\phi_1'/\partial x_\mu)(\partial\phi_1'/\partial x_\mu) + 2\mu^2\phi_2'^2] + 1/2(\partial\phi_2'/\partial x_\mu)(\partial\phi_2'/\partial x_\mu) + (\lambda/6)\chi\phi_1'(\phi_1'^2 + \phi_2'^2) - \lambda/24(\phi_1'^2 + \phi_2'^2)^2 \quad (4)$$

式子(4)中的有质量场 ϕ_1'对应于 χ 方向的振荡，而无质量场 ϕ_2'，对应于 ϕ 方向上的振荡。ϕ_2'的质量必须为零：因为相的对称性，能量不可能依赖于 α。戈德斯通猜测当连续对称性自发破缺时，一定存在零质量无旋粒子，即后来被称为戈德斯通玻色子或南部—戈德斯通玻色子。[21]

这个猜想的几个证明是由戈德斯通、萨拉姆和温伯格给出的 [289] (1962)。由于以下两点理由，这些在经典场有效的证明，在量子情形中却并不严密。第一，假定的守恒流中的因子不可对易。第二，对流的时间分量的积分可能不收敛，而流与玻色子场的对易关系对证明这个猜想是至关重要的。无论如何，在此论文发表后，作为连续对称性自发破缺结果的南部—戈德斯通无质量玻色子，已经成为某些场论模型的组成部分而被接受。

最初，戈德斯通对其结果态度不明朗。他写道："一种避开对称性的方法在基本粒子理论中当然被寄以很高期望。"然而，与自发对称性破缺相伴随的无质量玻色子的不存在，似乎使这种方法无用，除非用比戈德斯通简单模型"更复杂的方案"，否则无论如何不能消除这个困难(1961)。[22]

安德森解

把自发对称性破缺最终整合到规范理论框架又经过了两大步。贝克(Marshall Baker)和格拉肖在 1962 年，通过把自发对称性破缺的

思想扩展到非阿贝尔规范对称性的情行，而完成了第一步。在具有SU(3)对称性的模型中，他们发现不在拉格朗日函数中引入任何对称性破缺项，同样存在同位旋和超荷的低对称性解。安德森、希格斯(Peter Higgs)、基布尔和其他人不久就认识到把自发对称性破缺纳入有质量规范玻色子的规范理论中的重要性。

安德森(1963)迈出了第二步。他从施温格(1962a)的开创性论文出发；该论文论证了只要自耦合足够强，矢量规范玻色子并不一定要有零质量。安德森认识到施温格的意见具有潜在的重要性。安德森凭借对自发对称性破缺的超导案例的熟悉，发现超导性中的等离体子模式是一个

> 体现了施温格论点的实际例子，在某些情况下，杨—米尔斯类型的矢量玻色子不需要零质量。
>
> (1963)

正如我们前面注意到的，在超导性的BCS理论的最初解释中，没有零质量模。其中，由于能隙的原因，费米子质量是有限的，然而，无质量的戈德斯通玻色子通过与适当规范场即电磁场的相互作用，变成有限质量的等离体子。通过类比，安德森提出：

[290]

> 这种方法对南部类型的简并真空理论都是可行的，没有涉及零质量杨—米尔斯规范玻色子或零质量哥德斯通玻色子的困难。这两种类型的玻色子似乎都能“彼此抵消”，而只留下有限质量的玻色子。
>
> (同上)

安德森强调“给出规范场质量的……唯一机制是简并真空类型的理论”。而解决戈德斯通零质量困难的唯一机制，是具有无质量规范玻色子的杨—米尔斯理论(同上)。施温格—安德森机制(现在称为希格斯机制)，被许多学者在量子场论中认识和精致化，包括希格斯(1964a，b)，恩格勒特与布鲁(F. Englert and R. Brout，1964)，古拉尔尼克、哈根与基布尔(G. S. Guralnik，C. R. Hagen and T. W. B. Kibble，1964)在阿贝尔对称性的情形中所做的工作，以及基布尔在非阿贝尔对称性的情形中所做的工作，不久就成为规范理论概念框架不可或缺的组成部分。[23]

在20世纪50年代后期，物理学的各个分支取得了理论上的进展：博戈留波夫形式的超导BCS理论[24]，对其规范不变性的争论，以及对其非对称基态的认可；海森伯在其非线性旋量理论基础上建立的基本粒子复合模型，以及简并真空的思想；以及使得对南部—戈德斯通玻色子的认识变得容易的色散关系。由于这些进展，致使自发对称性破缺得以重新发现，并整合到规范理论的框架。自发对称性破缺最后整合到规范理论的框架，是作为曾困扰原初杨—米尔斯理论的短程核力作用机制问题的一个解而实现的。

自发对称性破缺的加入，已经深刻地改变了量子场论的概念基础。正如施韦伯曾经指出的：“1947—1952年期间强调协变方法的一个策略，是套用十分严格的相对论性不变性解释。”这种死板态度“导致了一个默认的假定，认为描述相对论性量子场论的真空态的海森伯态矢总是一种非简并矢量，并且这个矢量总是拥有拉格朗日函数的完全不变性”(1985)。南部证实了这个判断的准确性，他回想自己在自发对称性破缺方面的工作时说，“或许激励我作出贡献的，是对理论学家有关真空性质公理化的教条的挑战”(1989)，也就是采用

充盈真空观而不是空虚真空观。[25]

10.2 短程相互作用的机制(II)：渐近自由

在保持场论的规范不变性的前提下，另一种描述核力的短程行为的可能机制，也是迄今还没有很好地建立起来的机制，是在对重正化群分析的一般框架下，通过联合“渐近自由”(asymptotic freedom)和“色禁闭”(color confinement)的思想提出的。

[291]

总的语境是强子(强相互作用粒子：核子和介子)的夸克模型。在夸克模型中，强子被描述成为数不多的自旋为 1/2 的基本夸克的束缚态。这就使得是夸克而不是强子成为强核力的真正来源，而强子之间的总核力，即观测到的强核力，仅仅是更强大的夸克间作用力的残余力。

这就表明，如果在规范不变理论中用夸克和无质量规范玻色子(胶子)之间的有效耦合描述的夸克间的力，在短距离上趋于减小(渐近自由)，而在长距离上趋于发散(由胶子的无质量性造成)，以致强子的夸克和胶子作为荷(或色)状态都被禁闭在强子中，不可能被其他强子和它们的组分“看到”(色禁闭)，那么在强子之间将没有长程力，虽然强子之间的短程相互作用的精确机制仍付缺如。

从概念上讲，渐近自由是场的重正化群行为[像耦合常数这类物理参数依赖于所探测的(动量)大小]，与非阿贝尔规范场(或任何非线性场)的非线性(自耦合)相结合所产生的结果。正如我们在 8.8 节中提到的，威尔逊复苏了对重正化群的分析，而卡伦和西曼齐克进一步对它作了详尽阐述，其结果就是卡伦—西曼齐克方程组，该方程组简化了重正化群分析，并且能应用于各种物理情形，比如由威尔逊算符乘积展开所描述的情形。

卡伦和西曼齐克在渐近自由思想的发展过程中起了关键性的作

用。西曼齐克在完成了重正化方程的工作之后不久，试图建立一个渐近自由理论来解释比约肯标度无关性(见后)；他借助这些方程来研究在 ϕ^4 理论中耦合常数的巡行(running)行为*(1971)。他发现，如果耦合常数的符号这样选取以使得在相应的经典理论中能量有下限，那么耦合常数的大小将随着外部动量的增加而增加。西曼齐克证明，如果颠倒耦合常数的符号，那么相反的情况就会出现，并且在很高的外部动量下，将会有渐近自由。正如科尔曼和格罗斯(1973)所指出的，西曼齐克论证的困难在于，既然耦合常数的"错误"符号，只有当能量没有下限时才是可能的，那么，这样一个渐近自由系统将没有基态，因此也将是不稳定的。 [292]

受西曼齐克的激励，特霍夫特对渐近自由产生了浓厚兴趣。特霍夫特是第一个在规范理论的语境中建立渐近自由的物理学家(1972c)。在1972年6月的马赛会议上，在西曼齐克同一主题的报告(1972)之后，他以评论的方式汇报了自己的结果，认为诸如具有SU(2)群并且少于11种费米子的非阿贝尔规范理论，将是渐近自由的。这个报告非常著名并于1973年公布于众。

另外一条渐近自由的发展线索，开始于卡伦的工作，与强子现象有更为紧密的关系。在1968年的论文中，卡伦和格罗斯为深度非弹性电子—质子散射中可测量的结构函数提出了一个求和规则。随后比约肯注意到，这个求和规则意味着深度非弹性散射截面的标度无关性(1969)。用结构函数 W_i 表述这个截面，则比约肯标度无关性关系式为：

$$mW_1(\nu, q^2) \to F_1(x),$$
$$\nu W_2(\nu, q^2) \to F_2(x),$$

* "巡行耦合"是指耦合常数随动量标度而变化。——译者

$$\nu W_3(\nu,\ q^2) \to F_3(x) \tag{5}$$

(其中 m 是核子质量，$\nu = E - E'$，$E(E')$是轻子的初(终)能量，$q^2 = 4EE'\sin^2(\theta/2)$，$\theta$ 是轻子散射角，且 $0 \leqslant x(= q^2/2m\nu) \leqslant 1$)。这个关系式表明结构函数并不分别依赖于 ν 和 q^2，而只依赖于它们的比率 x，而 x 在 $q^2 \to \infty$ 时保持不变。

比约肯的预言不久就被斯坦福直线加速器中心(SLAC)的新实验所证实。[26]这些进展激发了提出各种理论方案的积极性(这些理论中的绝大多数是在光锥流代数和重正化群的框架下进行的)，它们非常有利于阐释强子的结构和动力学。

比约肯标度无关性的物理意义，不久为费恩曼在其强子的部分子模型中所揭示，此模型是深度非弹性散射的直观图景(1969)。按照部分子模型，强子可以被高能轻子束看到，就像是一个充满类点部分子的盒子；而在标度区域(高能区域或短程区域)中的非弹性轻子—核子散射截面，等于弹性轻子—部分子散射截面的非相干之和。[27]由于轻子—部分子散射的非相干性意味着轻子所看到的核子就像是一群可自由运动的组分，因此在场论语境中，比约肯标度无关性要求，部分子场的短程行为要用无相互作用的理论来描述。

[293] 卡伦(1970)和西曼齐克(1970)正是在其努力寻求比约肯标度无关性的前后连贯的理解的过程中，重新发现了重正化群方程，并且把它们表述成标度不变性反常的结果。[28]卡伦和西曼齐克随后把他们的方程应用到威尔逊算符乘积展开[29]，并一直推广到光锥(一个与深度非弹性散射有关的区域)，以便分析深度非弹性散射。[30]这些分析逐渐表明，一旦把相互作用引入到理论中，没有标度无关性也是可能的。

在此语境中会产生一个问题。如果在部分子中没有相互作用，强子应该容易被打碎成它们的部分子组分，然而为什么从来没有观测

到游离的部分子呢?

事实上，这个问题一直没有被很好地研究过，这是由于标度无关性仅涉及强子组分的短程行为，而打破强子却是一种长程现象。

然而，甚至在短程行为的范围内，标度无关性也曾受到过严峻的质疑，并且是建立在重正化群论证基础上的质疑。正如我们在8.8节提到的，按照威尔逊的观点，标度无关性只能在重正化群方程的不动点成立，这就需要在短程范围内，有效耦合要接近不动点的值。由于这些值一般是强耦合理论中的值，因此结果就会出现较大反常的标度无关行为，这与自由场论中的行为很不相同。因此，从威尔逊的观点看(见第8章的注释67)，人们可预期，非正则的标度无关性(伴随被吸收到场和流的反常维度的相互作用和重正化的效应)会在高能下出现，这种非正则标度无关性表示着重正化群方程的非平凡不动点。

事实上，认为在斯坦福直线加速器中心观测到的标度无关性不是真正的渐近自由现象的看法，已在理论学家中广泛传播。这种看法被如下事实进一步强化，即观测到的标度无关性现象甚至在相当低的动量转移情况下发生(所谓“早熟标度无关性”)。一直到1973年10月，盖尔曼和同事还推测，“可能有一个产生真正的标度无关性[威尔逊标度无关性]的高能修正”。[31]

在这种氛围下，格罗斯(丘的门徒，但后来转向盖尔曼的流代数纲领)在1972年底决定“证明定域场论不可能解释标度无关性的实验事实，因而不是一个描述强相互作用的适当框架”(Gross, 1992)。格罗斯采用两步策略。首先，他打算证明，有效耦合在短程上的消失，或者说渐近自由，为解释标度无关性所必需；其次，证明不可能存在渐近自由的场论。

[294] 事实上，帕里西(Giorgio Parisi，1973)不久就完成了第一步。他证明标度无关性只能用渐近自由理论解释，卡伦和格罗斯(1973)进一步把帕里西的思想扩展到除非阿贝尔规范理论之外的所有可重正化场论。卡伦—格罗斯论证是建立在威尔逊率先提出的标度量纲思想(见8.8节)之上的。首先，他们注意到，控制深度非弹性散射大小的复合算符有正则量纲。然后，他们证明了在重正化群的一个假定不动点处，正在减小组分算符的反常的标度量纲，必然包括正在减小的基础场的反常量纲。这反过来又意味着位于不动点的相互作用的基本场一定是自由的，因为通常只有自由的那些场才有正则标度量纲。他们总结说，只有当重正化群的一个设定的不动点位于耦合空间的原点处，即该理论必须是渐近自由的时，比约肯标度无关性才能得到解释。

格罗斯论证的第二步失败了。然而，这个失败导致对非阿贝尔规范理论渐近自由的一次重新发现，这一发现独立于特霍夫特，可以用来解释观测到的标度无关性，因此，具有讽刺意味的是，至少在非阿贝尔规范理论的情况下，确立了作为强相互作用描述框架的定域场论的适宜性。

起先，格罗斯关于没有任何理论会是渐近自由的这一断言听起来是合理的，因为在量子电动力学(即量子场论的原型)的情况下，有效电荷在短距离处增大，并且没有在特霍夫特的工作中存在的反例，这些格罗斯都是知道的。

作为电荷重正化的结果，量子电动力学中有效电荷会增大的原因易于理解。电荷重正化的根本机制是真空极化。按照狄拉克的观点，相对论性量子系统的真空是虚粒子(在量子电动力学情况下是虚电子—正电子对)介质。把一个电荷 e_0 置入真空就会使真空极化，

而具有虚电偶极子的极化真空则会屏蔽这个电荷，因此可观测电荷 e 也将不同于 e_0 而是成为 e_0/ε，这里的 ε 是介电常量。 由于 ε 的值与距离 r 有关，电荷的可观测值可用一个巡行有效耦合 $e(r)$ 描述，它决定了距离 r 处的力。 当 r 增加时，屏蔽介质增加，因此 $e(r)$ 随 r 的增加而减小，相应地随 r 的减小而增加。 于是，直接减去 $\log(e(r))$ 对 $\log(r)$ 的导数的 β 函数还是正的。

1973 年的春天，由于非阿贝尔规范理论吸引了物理学家的许多注意力——这主要是由韦尔特曼和特霍夫特的工作（见 10.3 节）引起的——因此，格罗斯在自己工作中不得不检查非阿贝尔理论的高能行为。 与此同时，科尔曼和学生波利策(H. David Politzer)出于不同的原因，也进行了类似的考察。 为了做到这一点，只要研究 β 函数在耦合常数空间的原点附近的行为就够了。 在非阿贝尔规范理论中对 β 函数进行的计算[32]表明，与格罗斯的期望相反，理论是渐近自由的。 [295]

导致非阿贝尔（色）规范理论（见 11.1 节）跟量子电动力学这种阿贝尔规范理论表现行为不同的原因，在于非阿贝尔规范场的非线性或自耦合。 也就是说，在于这样一个事实：除了携带色荷的夸克外，规范玻色子（胶子）也携带色荷。 就虚夸克—反夸克对而言，发生在阿贝尔情形中的屏蔽，同样也发生在非阿贝尔理论中。 然而，携带色荷的胶子仍然对真空极化作了额外的贡献，主要倾向是强化而不是中和规范荷。[33]详细计算表明，如果夸克三重态的数目少于 17，那么源于虚胶子的反屏蔽就会胜过虚夸克—反夸克对产生的屏蔽，并且系统是渐近自由的。

科尔曼和格罗斯(1973)也证明，任何由汤川式的、标量的或阿贝尔规范相互作用构成的可重正化场论，都不可能是渐近自由的，而任

何渐近自由的可重正化场论一定包含非阿贝尔规范场。 这个推论有效地限制了描述强相互作用的动力学系统的选择。

根据他们对非阿贝尔理论中有效耦合的特殊行为(在短程时，它趋于衰减；在长程时，它变得很强)的理解，以及由对标度无关性的观测所提供的信息，格罗斯和维尔切克(Gross and Wilczek, 1973b)讨论了把一个没有破缺对称性的非阿贝尔规范理论作为描述强相互作用的框架的可能性。 他们指出，杨—米尔斯理论的最初困难，即要有传递长程相互作用的强耦合无质量矢量介子，在渐近自由理论中消失了。

色禁闭

他们认为，色禁闭的原因如下。 强子(夸克和胶子)的色荷组分的长程行为，在这种情形中可能是由理论的强耦合极限所决定的，并[296]且这些组分是被禁闭在强子内。 因此在长程范围内，除了色中性粒子(色单态或强子)以外，所有的色荷态都可能被抑制而不能从强子之外“看见”。 并且这与强子之间不存在长程相互作用是一致的。

应当注意的是，虽然人们用这些术语能富有启发性地理解禁闭思想，并且也作过许多努力[34]，但是从来没有出现过关于禁闭的精确证明。

10.3 可重正化性

任何一个建立在非阿贝尔规范理论框架下的模型(不论有没有有质量规范玻色子，是不是渐近自由的)，除非被证明是可重正化的，否则都没有任何预测能力。 因此规范理论要成为一个成功的研究纲领，人们不得不证明它的可重正化性。 当杨振宁和米尔斯提出他们关于强相互作用的非阿贝尔规范不变理论时，其可重正化性不过是假

设的。这个假设仅仅建立在方次计数论证（power-counting argument）* 上：在量子化和导出费恩曼图的规则之后，他们发现，虽然费恩曼图中的基本顶点有三种类型（因为带荷规范玻色子的自耦合）而不是像在量子电动力学中只有一种类型，但"这个'原始发散'在数目上还是有限的"（1954b）。因此，按照戴森的重正化纲领，人们没有理由担心其可重正化性。一直到 1962 年费恩曼开始考察这个问题（见下文），这个信念才受到挑战。事实上，即使在费恩曼的论著（和受此激发出版的其他论著）出版之后，许多物理学家还是持有这种观念，比如温伯格（1967b）和萨拉姆（1968）。

随着 1957 年弱相互作用中宇称不守恒的发现，以及 1958 年弱相互作用中矢量和轴矢耦合理论的建立，出现了许多关于传递弱相互作用的是中间矢量介子 W 和 Z 的推测（见 8.6 节和 9.3 节）。人们试图把矢量介子等同于非阿贝尔规范理论的规范玻色子。由于弱相互作用的短程特征要求带荷中间介子是有质量的，因此这种等同的结果就是所谓的有质量杨—米尔斯理论。有质量杨—米尔斯理论的出现开创了两条研究路线。除了寻求一个"软质量"机制（通过这一机制，规范玻色子获得质量而理论却保持规范不变）外，更多的努力是试图证明有质量带荷矢量理论的可重正化性。

[297] 研究非阿贝尔规范理论的可重正化性的漫长过程始于布卢德曼的评论：把非阿贝尔规范玻色子解释成实粒子的一个困难是，带荷矢量玻色子具有奇异的、不可重正化的电磁相互作用。布卢德曼的评论是基于早年就已很好地建立起来的介子电动力学的一个结论。[35]

然而，众所周知，带电矢量介子的非可重正化特征甚至在处理电

---

* "方次计数"是估算费恩曼积分发散程度的一种简便方法。——译者

磁相互作用之外的其他相互作用时也存在。比如，在讨论矢量介子与核子的矢量相互作用时，马修斯(Matthews，1949)注意到中性矢量介子和带电矢量介子的相互作用之间的一个重大区别：前者被证明可重正化，而后者情况却并非如此。马修斯的发现随后被格劳泽(Glauser，1953)、梅泽(1952)以及梅泽和龟渊(1951)精致化。通过强调源流的一种假定的守恒，马修斯含蓄地而格劳泽以及梅泽和龟渊却明确地认为，只有在中性矢量介子的情况下才存在守恒，如果矢量介子带电，就不存在守恒。应当注意的是，在所讨论的这个发展阶段，荷流的不守恒被取作破坏带电矢量介子理论的可重正化性的附加项的主要来源。

布卢德曼的评论向施温格的规范纲领提出了一个严重问题。根据这个纲领，弱相互作用的所有现象都是由具有普遍规范耦合的矢量介子决定的，这种介子是不稳定的、有质量的，并且是有荷的。对布卢德曼的评论立即作出回应的是格拉肖(1959)(施温格的门生)，随后作出反应的有萨拉姆、梅泽、龟渊、沃德、科马尔(Komar)和其他许多人。[36]人们得出了与带电矢量介子理论的可重正化性相冲突的断言。但这些回应在理解真实情况方面都没有进展，直到李政道和杨振宁论文(1962)的发表之后才有了真正的进展。

为了系统地研究这个问题，李政道和杨振宁从有质量带电矢量介子理论导出费恩曼图规则入手。他们发现存在一些发散的并且是相对论性非协变的顶点。通过引入一个极限程序(对拉格朗日函数添加一项$[-\xi(\partial_\mu^*\phi_\mu^*)(\partial_\nu\phi_\nu)]$，这个项依赖于正参数$\xi\rightarrow 0$的过程)，费恩曼规则变成协变的，不过这个理论仍然按一种不可重正化方式发散。为了弥补这一点，他们引入了一个负度规，使参数ξ起到调节子的作用。最终理论在$\xi>0$时既是协变的也是可重正化的。取得这一成

功的代价是，负度规的引入破坏了理论的幺正性。

李政道和杨振宁的体系缺乏幺正性，在物理上的表现就是出现非物理标量介子。 由于幺正性的物理意义是概率的守恒，它本身是任何有意义的物理理论必须满足的少数几条基本要求之一，所以摆在李政道和杨振宁面前的困难，是如何改变这些非物理的自由度而实现幺正性，这个困难在重正化非阿贝尔规范理论时成为中心问题，也是随后几年追求的目标。 [298]

非阿贝尔规范理论的重正化的下一个里程碑是费恩曼的工作(1963)。[37]在形式上，费恩曼的主题是引力。 但为了比较和说明，他也讨论了另一个非线性理论，即杨—米尔斯理论，其中相互作用也是用无质量粒子传递的。 费恩曼讨论的中心议题是导出一套图形规则(尤其是圈图规则)，这些规则既是幺正的又是协变的，以便能够进一步系统地研究这些规则的可重正化性。

费恩曼对幺正性的讨论，是通过树图定理彼此联结起来的树(trees)和圈(loops)这些术语进行的。 根据树图定理，封闭的圈图由在实际区域上的和在质壳上的相应的树图共同构成，它们是通过打开圈中的一个玻色子线得到的。 另外，费恩曼在具体计算一个圈(每一个树图中无质量玻色子线都是开口的)时，只考虑到玻色子的实横向自由度，以保证规范不变性。 费恩曼在树图定理背后的物理学动机是表达物理量的强烈愿望，这些物理量在量子场论中通常用虚场(圈图)的方法定义，在真实物理过程中则用实际可测的量定义。

按照费恩曼的意思，幺正性关系是在封闭圈图和相应的物理树集之间的一种关联，但除此之外还有许多其他关联。 事实上，他发现从通常路径积分公式导出的圈图规则不是幺正的。 费恩曼用把一个圈图断开成一个树集的方法，清楚地揭示了圈图的幺正性规则失败的

根源。在一个协变数学公式(这是重正化所必需的)中,圈图中的一条虚线,除了表示横向极化之外,还表示长度或标量自由度。一旦它被断开,就变成一条外线,代表一个自由粒子,它就仅仅表示(物理上的)横向自由度。因此,为了保持洛伦兹协变性(并把圈图与树图相匹配),不得不把有些非物理的额外自由度考虑进去。这样费恩曼被迫面对一个额外粒子,一个"荒唐"或"虚幻"的粒子,它本来属于圈图中的传播子,并且是必须被整合掉的,但现在变成了自由粒子。因此幺正性被破坏,并且费恩曼认识到"有些事情[是]在根本上错了"(1963)。错误出在两个基本的物理学原则——洛伦兹协变性和幺正性——之间的正面冲突。

[299]

为了恢复幺正性,费曼提出要从圈图中的玻色子内线"减去一些东西"。这个富有远见的提议是对该课题概念发展的第一个清楚认识。这个课题就是不得不引入某种鬼圈(ghost loop),以便能够得到非阿贝尔规范理论的连贯的形式体系。这个洞见立刻被出席费恩曼演讲的一位眼光敏锐的物理学家抓住,并要求费恩曼,"如果你想直接用圈图来重正化任何东西,就要更详细一点地给出虚幻粒子的结构和性质"(同上)。这位敏锐的物理学家就是德威特(Bryce DeWitt),他自己也正在对此课题的发展作出实质性贡献(见下文)。

除了由组合所致的对图形规则幺正性的证实[38],费恩曼也讨论了有质量问题,并试图通过取零质量极限得到无质量杨—米尔斯理论。这个动机是双重的。第一,他想回避无质量理论的红外问题,因为规范玻色子中的自耦使这个问题比在量子电动力学中的情况还糟糕。第二,不可能写下无质量矢量粒子的明显协变的传播子,这种粒子只有两种极化状态,因为,用协变方法从质壳中扩展出去,而无质量矢量粒子会获得第三种极化状态。这就是我们刚才提到的协变性和幺

正性在无质量矢量理论情况下的同一种冲突。

在这种语境下，有两点值得注意。第一，取零质量极限可能得到无质量理论这样一种假定，最终证明是一个错误(见下文)，因为有质量矢量玻色子的纵向极化在零质量极限时没有退耦。第二，费恩曼试图找到“一种重新表示圈图的方法，当 $\mu^2 \neq 0$ 时，有一个不同[于方程(10)，即 $(g^\mu_\nu - q^\mu q_\nu/\mu^2)/(q^2 - \mu^2)$] [39]的传播子的新形式……在这样一个形式下能取 μ^2 趋于零的极限”。为了这个目的，他引入了一系列被随后的研究者所沿用的处理方法。无论如何，这是首次明确提出寻找不同的规范，以便可能改变有质量矢量粒子的传播子的费恩曼规则。这个提议的结果即所谓的费恩曼规范，其中矢量传播子取 $[g^\mu_\nu/(q^2 - \mu^2)]$ 形式。费恩曼规范是可重正化规范的一种形式，其中有质量杨—米尔斯理论最后证明是可重正化的，因为导致所有问题的 $q^\mu q_\nu/\mu^2$ 项没有了。

受费恩曼工作的激发，德威特(1964，1967a，b，c)思考了规范选择问题和与之相关的鬼粒子(费恩曼提出的“虚幻粒子”)问题。虽然幺正性问题没有讨论清楚，但在引力情况下正确的费恩曼规则却按费恩曼规范建立了起来。另外，德威特规则中的鬼部分表述成包含复标量场的定域拉格朗日函数形式，这部分服从费米统计。[40] [300]

在规范对称性自发破缺的情况下，恩格勒特、布鲁和西里(F. Englert，R. Brout and M. Thiry，1966)也在朗道规范中导出了许多费恩曼规则，其中矢量传播子取 $[(g^\mu_\nu - q^\mu q_\nu/q^2)/(q^2 - \mu^2)]$ 的形式，并且找到相应的鬼。在矢量传播子形式的基础上，他们提出，理论是可重正化的。

法捷耶夫和波波夫(L. D. Faddeev and V. N. Popov，1967)从费恩曼的工作出发，并且使用路径积分形式，导出了朗道规范中的费恩曼

规则。在他们的工作中，鬼是直接产生的。但是他们的方法不能有效地用来讨论一般规范，也不能用来理解之前建立的费恩曼规范的规则。尤其是，他们引入的鬼，跟德威特的鬼不一样，没有方位，并且只有在朗道规范中才能获得简单的顶点结构。由于在他们的工作中选择规则和鬼规则都不是按拉格朗日函数的定域场表述的，因此拉格朗日函数和费恩曼规则之间的关联也不甚清楚。这项工作的另一个缺点是幺正性问题还没有交待。由于如上所述鬼的根源在于幺正性的要求，而在法捷耶夫和波波夫引入鬼作为规范不变性破坏的一个补偿的过程中，幺正性并非是一个无足轻重的问题，因此这一缺点在他们的推理中留下了一大缺陷。

德威特以及法捷耶夫和波波夫的结论，曾经由曼德尔施塔姆(1968a，b)用场论的完全规范不变性形式重新推导出来。因此在1968年夏天，无质量非阿贝尔规范理论的费恩曼规则已经建立起来，并且至少就某些具体的规范选择而言，它们在方次计数意义上是可重正化理论的，以致这些规范选择叫做可重正化规范。然而，无穷大是不是能够按微扰方法一步一步消除，在规范不变方法中还是不清楚的。另外，幺正性问题尚未解决，并且红外问题仍然困扰着物理学家。

在这个关键时刻，荷兰乌得勒支的韦尔特曼(1968a，b)对非阿贝尔规范理论的可重正化难题发起了英勇的攻击。韦尔特曼对当时的观点知道得很清楚，当时人们认为带电有质量矢量玻色子理论，作为弱相互作用的V-A理论的结果的复活，是毫无希望地发散的。流代[301]数的成功被当作场论的规范不变性的一个结果，然而在韦尔特曼看来，这种成功，特别是阿德勒—韦斯伯格关系的成功，只是有质量矢量玻色子的非阿贝尔规范理论的可重正化性的实验证据。于是，韦

尔特曼开始仔细研究以证实他的理解。

就现象而论，把矢量介子看成是有质量的这一点很自然，因为矢量介子要负责短程弱相互作用的传递就必须是有质量的。理论上讲，有质量理论能绕过红外问题，因为“零质量理论包含可怕的红外发散”，也正是因为这一点而无法进行具有入射和出射软玻色子的S矩阵研究。追随费恩曼，韦尔特曼也错误地认为无质量理论可以从有质量理论取零质量极限而得到。

有质量理论显然是不可重正化的，主要原因是在矢量场传播子 $(\delta_{\mu\nu}+k_\mu k_\nu/M^2)/(k^2+M^2-\mathrm{i}\varepsilon)$ 中含有 $k_\mu k_\nu/M^2$ 这一项。韦尔特曼熟悉费恩曼、德威特等人的早期工作，他注意到，“众所周知，如果人们对这一项进行规范变换，它就会被修正”，因为随着规范发生改变，费恩曼规则也可能改变。他的目标是找到改变规范的方法，以使费恩曼规则可按如下方式改变，即矢量传播子可取 $\delta_{\mu\nu}/(k^2+M^2-\mathrm{i}\varepsilon)$ 的形式[41]，并且这个理论至少在方次计数意义上是可重正化的。

韦尔特曼称他所发明的这种方法为自由场方法，而称关键的变换为贝尔—特赖曼变换，虽然不论是贝尔还是特赖曼对此都没有贡献。这种方法的基本思想是引入一个与矢量玻色子没有相互作用的自由标量场。通过用矢量场和标量场的联合（$W^\mu+\partial^\mu\phi/M$）来代替矢量场（W^μ），再加上以使标量场仍然是自由场的方法来增加新的顶点，韦尔特曼发现，虽然新理论的费恩曼规则不同（矢量场的传播子被联合的情况所代替），但S矩阵及相关物理学仍保持一样。他还发现，这种联合也可能选择成这样一种方式，其传播子采用的形式为

$$(\delta_{\mu\nu}+k_\mu k_\nu/M^2)/(k^2+M^2-\mathrm{i}\varepsilon)-k_\mu k_\nu/M^2/(k^2+M^2-\mathrm{i}\varepsilon)=\delta_{\mu\nu}/(k^2+M^2-\mathrm{i}\varepsilon),\qquad(6)$$

它可能导致较少发散的费恩曼规则。这一成果所付出的代价是引入了涉及鬼的新顶点。[42]

韦尔特曼的自由场方法是用图的语言发展起来的，并受丘的S矩阵哲学的激发。后来博尔韦尔(D. Boulware)在1970年论文中重新用路径积分形式对其进行改造，弗拉德金和秋京(E. S. Fradkin and I. V. Tyutin，1970)也做了同样的工作，两人在拉格朗日函数中引入一项(从而有可能引入规范的一般选择)，并且建立起一个程序，据此可得到一般规范的费恩曼规则。

[302]

有了改变规范和改变费恩曼规则的强有力方法，韦尔特曼得到一个惊人结果，也就是具有一明确质量项的杨—米尔斯理论，可以证明，这种理论是关于方次计数(即直到自能的二次项，三点顶点的线性项，和四点函数的对数项)单圈可重正化的。在这一突破之前，涉及规范理论的可重正化性的断言和说法仅仅是猜测。这一突破打破了认为带电有质量矢量玻色子理论是没有希望地不可重正化的原有看法，并激发了关于这个主题的进一步研究。

在概念上和心理上，韦尔特曼的论文(1968b)都极其重要，因为它使物理学家确信，通过改变规范，一套可重正化的费恩曼规则是可得到的。不过，在物理上还有许多缺口有待填平。为了恰当地重正化一个具有规范对称性的理论，人们不得不按照这种对称性对其规则化。在韦尔特曼(1968b)的情况中，理论没有对称性，因为质量项确实毁坏了假定的对称性。另外，他还没有规范不变的规则化方案。由于这些原因，该理论即使在单圈图情况下也不是完整意义上可重正化的。

这种情况在二圈图或多圈图时更糟糕。人们不久就发现，(通过在分母中加上质量来修正无质量规则而得到的)在单圈图层次上有效

的费恩曼规则，在二圈图上就明显违背了幺正性并导致不可重正化发散。[43]博尔韦尔在其路径积分处理中也表明，在二圈图及多圈图层次上会产生不可重正化顶点。因为在无质量理论中没有这种顶点，因此估计他们可能略去一二种方法。不过，斯拉夫诺夫和法捷耶夫(A. Slavnov and Faddeev，1970)明确了，有质量理论的零质量极限，在单圈图层次上，或者如范达姆和韦尔特曼(H. van Dam and Veltman，1970)所证明的对引力而言在树图层次上，已经不同于无质量理论。

因此更大的进展，有赖于对二圈图或多圈图的精确费恩曼规则是什么、以及对如何证明由这样的规则描述的理论的幺正性的恰当理解。一般来说，非阿贝尔规范理论情况下的幺正性问题比在其他情况下更复杂。可重正化规范中的费恩曼规则显然不是幺正的。存在非物理的自由度或鬼，这是用来补偿在法捷耶夫—波波夫公式中由规范固定项引入的规范依赖性。另外，在其协变形式的内部传播中，还存在规范场的非物理极化状态。只有当这些非物理的自由度不起作用时，一个理论才是幺正的。结果就会出现如下情况，即如果某些关系式或涉及规范场或鬼场的格林函数之间的沃德恒等式有效，那么这些作用将不存在。也就是说，沃德恒等式将确保在两种类型的非物理自由度之间的作用会严格抵消。因此为了满足幺正性条件，沃德恒等式的建立是关键。 [303]

就起源而论，沃德恒等式是沃德(1950)导出的，并且由弗拉德金(1956)和高桥(Takahashi，1957)扩展为量子电动力学中规范不变性的一个结果：

$$k_\mu \Gamma_\mu(q, p) = S_F^{-1}(q) - S_F^{-1}(p), \; k = q - p, \qquad (7)$$

式中 Γ_μ 和 S_F 是不可约化的顶点和各自的自能部分。这些恒等式说的是，如果一条或多条外线的极化矢量 $e_\mu(k)$ 被 k_μ 替代，而其他所有线都保持在质壳上并且由实际极化矢量提供，那么结果为零。

在非阿贝尔规范理论的情形下，情况复杂得多了。除了那些与出现在量子电动力学中的相似情况之外，还存在只涉及非阿贝尔规范场以及与鬼有关的一组恒等式。而且，由于这些恒等式与不可约化的顶点及自能部分无直接关系，它们的直接物理意义也就比在量子电动力学中的还少。不过，韦尔特曼(1970)通过使用施温格的源方法和贝尔—特赖曼变换，成功地为有质量非阿贝尔规范理论导出了广义沃德恒等式。[44]这些恒等式跟韦尔特曼(1963a, b)早年得到的截止方程组(cutting equations)一起，都能用于证明包含有鬼圈的非阿贝尔规范理论的幺正性。它们在实现真正的重正化过程中起了关键作用，因为重正化理论一定要证明是幺正的。

因此，到1970年夏天，韦尔特曼已经知道了如何从费恩曼图导出费恩曼规则，如何通过贝尔—特赖曼变换改变规范，从而得到不同的(幺正的或可重正化的)规则，以及如何通过使用广义沃德恒等式和他多年发展起来的复杂组合方法证明幺正性。所有这些都是重正化非阿贝尔规范理论的关键部分。然而，韦尔特曼实际达到的仅仅是对有质量矢量理论的单圈可重正化性质的一个说明，这也仅仅是建立在硬性操作和朴素的方次计数论证基础之上，而没有严格证明。对[304]他而言，一旦超出单圈图情形，这个理论就是不可重正化的。在韦尔特曼的计划中一个更严重的缺陷，是缺乏恰当的规则化程序(regularization procedure)。要证明所有抵消都能按一致的方法定义而不违背规范对称性，就必须有一个规范不变的调节子方法。因此，当正在他的指导下研究高能物理学的学生特霍夫特表现出对杨—

米尔斯理论的兴趣时，韦尔特曼对他提出一个问题，即寻找一个用于杨—米尔斯理论的好的调节子方法。这一提议不久就导致了最终突破。

在特霍夫特的论文(1971a)中，他对单圈图提出了一个新的截止方案，在此方案中，他引入了第五维而使理论的规范对称性得以保持，并把这种方法用于无质量杨—米尔斯理论。他优美地推广了法捷耶夫和波波夫(1967)关于一般规范固定和鬼的生成的方法，以及弗拉德金和秋京(1970)在一般规范中费恩曼规则的导出方法，并且提出了包括按定域非重正化拉格朗日函数方式表示的一大类规范中的定向鬼的费恩曼规则[45]，从而打开了通向完全重正化纲领之路。通过使用韦尔特曼的组合方法，特霍夫特能够建立起广义沃德恒等式，并且借助于它们达到幺正性。不过，特霍夫特关于无质量杨—米尔斯理论的可重正化性的证明只对单圈图有效，原因是新的截止方法只作用在单圈图上。

在1971年，无质量杨—米尔斯理论被看作一个非实在的模型，而且只有有质量理论，借助于带电矢量玻色子，才能被看作物理相关的。由于在韦尔特曼先前的研究(1967, 1970)中，有质量理论明显是不可重正化的，因此他十分希望扩大特霍夫特的成功，并希望得到有质量带电矢量玻色子的可重正化理论。有从韦尔特曼那里学到的方法作武器，并在本杰明·李(Benjamin Lee)关于 σ 模型的重正化工作的激励下，特霍夫特自信地承诺会给出一个解答。

特霍夫特的基本思想是把自发对称性破缺引入杨—米尔斯理论。这个思想受以下进展所推动。第一，矢量玻色子通过希格斯型机制能得到质量。第二，无质量理论的红外困难能因此而绕过去。第三，也是最重要的，无质量理论的可重正化性得以保留，因为正如本

杰明·李和其他人关于 σ 模型的工作所证明的[46]，理论的可重正化性没有被自发对称性破缺损坏。一旦实现了这一步，用于无质量理论的技巧和由韦尔特曼为有质量理论的单圈图可重正化性而发展的那些技巧，就可以用来证明有质量杨—米尔斯理论的可重正化性('t Hooft，1971b)。

[305] 这篇论文立即被誉为“将以最深远的方式改变我们对规范场论的思维方式”[47]，并且与特霍夫特的另一篇论文(1971a)一起，被普遍认为是一个转折点，这一点不仅是对规范理论的命运，也是对一般量子电动力学的命运和许多建立在其上的模型而言的，因此它根本上改变了基础物理学，包括粒子物理学和宇宙学的进程。

虽然这个评价基本上是正确的，不过，要完成重正化纲领，还有两个缺口不得不填平。首先是规范不变的规则化程序，不仅在单圈图层次上，而且是对微扰理论中任何有限阶，都要被看作有效的。这个程序之所以特别重要，是因为随后必须作抵消处理。这个缺口是特霍夫特和韦尔特曼(1972a)填平的，在这篇论文中，他们提出了称为维数正规化(dimensional regularization)的系统程序，它建立在特霍夫特规范不变的截止的扩展(在其中引入了时空的第五维)基础上，通过让时空维度取非整数值、并且在环积分进行之前就确定其连续性。[48]

第二，必须一个接一个牢固地建立起沃德恒等式、幺正性和重正化。特霍夫特和韦尔特曼(1972b)提出了一种形式上的沃德恒等式的联合推导，并借用它们进行非阿贝尔规范理论的幺正性及可重正化性的证明，因此为规范理论纲领的进一步发展打下了坚实基础。

韦尔特曼—特霍夫特纲领的发展彻底清楚地说明，保证规范不变性对任何矢量玻色子理论的可重正化性的证明都是至关重要的。这

有两个原因。第一，建立一套可重正化费恩曼规则的幺正性，要求有一组沃德恒等式作为规范不变性的结果。第二，可重正化规范中的规则，仅当理论是规范不变时等价于幺正规范中的规则。这个结论严重制约了一旦涉及玻色子场时量子场论的进一步发展。[49]

10.4 整体特征

传统上，场论研究物理世界的局部行为。在规范理论的情况下，局部相互作用直接决定了规范势的局域特征，也决定了守恒定律。还有，随着规范理论框架的兴起，另外一些神秘的物理现象在此框架内也逐渐变得可理解，更为重要的是借此揭示了规范势的某些整体特征，这里的整体性是指规范势的所有拓扑等价(可变形的)状态 [306] (或等价类)有共同的拓扑行为，而不同的类则有不同的行为。

质子和电子的电荷绝对值相等就是这样一种现象。在实验上，它们在很高的精确度上都相等。在理论上，狄拉克(1931)试图通过提出磁单极的存在来理解这种相等性，这意味着电荷的量子化。虽然至今还没有探测到磁单极[50]，但是这个主张确实触动了规范不变场论的理论结构的核心。规范理论在这方面的进一步探索已经揭示了规范势的整体特征，并使物理参量的量子化变得非常清楚，而不管是否隐含着磁单极的存在。杨振宁(1970)所作的另一努力是按照施温格的方法，试图把电荷的量子化和规范群的紧致性联系起来，后者是这种群的一种整体特征。[51]

属于这类的另一现象是著名的阿哈罗诺夫—玻姆(AB)效应。在由阿哈罗诺夫和玻姆提出、并且由钱伯斯(R. G. Chambers)完成的一个实验中，在电磁场强度 $F_{\mu\nu}$ 处处为零的螺线管外的双连通区域中出现的现象(电子束波函数中附加相位的移动引起干涉条纹的移动)，最

终依赖于 A_μ 沿一个不可收缩环路的积分，$\alpha = e/hc\oint A_\mu \mathrm{d}x^\mu$（或更准确地说，依赖于相位因子 $e^{i\alpha}$ 而不是相位 α）。

AB 效应的含义很多，其中最重要的有两方面。第一，它确立了规范势 A_μ 的物理实在性。第二，它揭示了电磁场的非局域或者说整体特征。

传统观点认为 $F_{\mu\nu}$ 是物理实在，具有可观测效应，并给出了电磁场固有的、完备的描述。相反观点是把势 A_μ 仅仅当作辅助量并且是虚设的，没有物理实在性，因为它被认为是任意的而不能产生任何可观测效应。AB 效应已清楚地表明，在量子理论中 $F_{\mu\nu}$ 自身不能完备地描述电子波函数的所有电磁效应。由于有些效应只能用 A_μ 独自描述，同时 $F_{\mu\nu}$ 自身总能用 A_μ 表示，因此，可合理地假定势 A_μ 不仅在它能产生可观测效应的意义上具有物理实在性，而且必须被看成比电磁场强度更基本。既然电磁场有实体性，势也一定是实体的，这样才能使得场从势中获得其实体性。这是 AB 效应对规范场纲领的本体论认同的一个重要含义。

[307]

第二，在场强不存在的区域观测到 AB 效应，也说明了由势所产生的电子波函数的相位的相对变化是物理上可观测的。然而，正如我现在就要讨论的，这种变化与其说是由任何具体势跟电子的任何局部相互作用所产生的，毋宁说它是由势的某种整体性质决定的，这种整体性质是由纯规范函数指定的，并对特定的规范群有唯一性。那就是说，根据其本体论基础，规范势在其“真空”中或者说在“纯规范”构形中是唯一的。因此最初反对规范势的实在性的意见——认为它有任意性——在此看来已没有说服力。

在更大的理论语境中详尽叙述第二点这个我们这一节的主题之

前，先让我提一个物理上更明显的现象；如果不考虑到规范理论的整体特征，这个现象就会很难理解。这就是1969年发现的阿德勒—贝尔—贾基夫(ABJ)反常(Adler-Bell-Jackiw anomalies)(见8.7节)。与非阿贝尔规范理论的有效拉格朗日函数中的这些反常相关联的是一个附加项，该项与 $\boldsymbol{P}$ 成正比，而 $\boldsymbol{P} = -(1/16\pi^2)\mathrm{tr}\boldsymbol{F}^{*\mu\nu}\boldsymbol{F}_{\mu\nu} = \partial_\mu \boldsymbol{C}^\mu[\boldsymbol{C}^\mu = -(1/16\pi^2)\varepsilon^{\mu\alpha\beta\gamma}\mathrm{tr}(\boldsymbol{F}_{\alpha\beta}\boldsymbol{A}_\gamma - 2/3\boldsymbol{A}_\alpha\boldsymbol{A}_\beta\boldsymbol{A}_\gamma)]$(参阅8.7节中的方程(25))。这个规范不变项对运动方程没有作用，因为它对规范不变量 $\boldsymbol{C}$ 总体上是发散的。这意味着：第一，反常引起的规范对称性的破缺有量子根源，因为经典动力学是可以完全由运动方程描述的。第二，它提示了反常与规范势的长程行为有关联。更准确地讲，它提示了反常所破坏的是在某些有限规范变换下的对称性，而非无穷小规范变换下的对称性。

所有这些现象的共同点是有限规范变换的存在，这些有限规范变换(不像无穷小规范变换)作为一种平行位移，都用路径相关的或不可积的相位因子表示。[52]在狄拉克磁单极情形中，规范变换不可积这一点很容易理解。正是磁单极的存在造成规范势中的奇点，这使得奇点周围的势的环积分不可定义。在AB效应中，螺线管的存在起着狄拉克磁单极同样的作用：相因子中的环积分不得不沿一个不可收缩环进行。在反常的情况中，上面提到的附加项在无穷小变换下是不变的，但是在某些边界条件下，在非阿贝尔理论有限变换下，要增加 [308] 一个附加项 ($\boldsymbol{A} \rightarrow S_U\boldsymbol{A}S_U^+ = U^{-1}(\boldsymbol{r})\boldsymbol{A}U(\boldsymbol{r}) - U^{-1}(\boldsymbol{r})\nabla U(\boldsymbol{r})$)，这是由规范势的长程性质 $A_\mu(\boldsymbol{r}\rightarrow\infty) \rightarrow U^{-1}\partial_\mu U$ 决定的：

$$\omega(U) = 1/24\pi^2\int \mathrm{d}\boldsymbol{r}\varepsilon^{ijk}\,\mathrm{tr}(U^{-1}(\boldsymbol{r})\partial_i U(\boldsymbol{r})U^{-1}(\boldsymbol{r}) \cdot \partial_j U(\boldsymbol{r})U^{-1}(\boldsymbol{r})\partial_k U(\boldsymbol{r})), \tag{7}$$

其中 $U(\boldsymbol{r})$ 是有长程渐近线 I（或 $-I$）的纯粹规范函数。在此情况下，有限和无穷小规范变换的区别在于涉及规范函数 $U(\boldsymbol{r})$ 的项。因此更进一步地考察规范函数看来是值得的。

事实上，对规范理论的整体特征的一种深入理解，就源自对规范函数 $U(\boldsymbol{r})$ 的如下考察：由于有无穷远点的三维空间流形，在拓扑学上视为等效于 S^3，即用三个角标记的四维欧氏球的表面，因此在数学上，U 函数提供了从 S^3 到规范群流形的映射。从附录 2 我们知道，这种映射属于分离同伦类，用整数标记；属于不同类的规范函数不能彼此连续变形；而且只有那些同伦平凡群（其绕数等于零）才能变形到同一。也就是说，Π_3（非阿贝尔规范群）$= \Pi_3(S^3) = Z$。而且，在非阿贝尔理论的有限变换（$\boldsymbol{A} \rightarrow S_U \boldsymbol{A} S_U^{\dagger} = U^{-1}(\boldsymbol{r})\boldsymbol{A}U(\boldsymbol{r}) - U^{-1}(\boldsymbol{r})\nabla U(\boldsymbol{r})$）下，由韦斯祖米诺项（Wesszumino term，表示反常效应）获得的附加项 $\omega(U)$，正如在附录 2 中可看到的，就是有限规范变换绕数的分析表达式。

有限规范变换的概念内涵丰富，它用绕数标记的性质由整体规范变换 $U_\infty (= \lim_{r\rightarrow\infty} U(\boldsymbol{r}))$ 唯一地决定。首先，在任何具有手征源流的实际非阿贝尔规范理论中，其规范变换是不平凡的并且手征反常是不可避免的；规范不变性在量子层次上被破坏，并且可重正化性也受损，除非对源流作些处理以使反常项确定的附加贡献彼此抵消，从而恢复规范不变性。抵消反常的要求已成为建立模型的一项严格约束条件（见 11.1 节）。

第二，当三维规范变换不是平凡的（比如非阿贝尔规范理论的情况，但不限于此）时，如果作用 $\boldsymbol{I}$ 相对于小的规范变换而不是大的规范变换是不变的，那么在某些情形下的规范不变性要求就意味着物理参量的量子化。为了理解这一点，我们只要回忆一下，在路径积分

[309]

的形式中，理论由 $\exp(\mathrm{i}\boldsymbol{I}/h)$ 决定，相对于有限规范变换，理论的规范不变性能够通过如下要求达到，这个要求就是由有限规范变换引起的作用量的改变 $\Delta\boldsymbol{I}$ 等于 $2n\pi h$，其中 n 是整数。在狄拉克的点磁单极情形中，由于存在磁单极引起的奇点以及 $\Delta\boldsymbol{I}=4\pi eg$，$U(1)$ 规范变换是非平凡的，使得规范不变性的要求等同于要求 $4\pi eg/c=2n\pi h$ 或 $2ge/hc=n$ 这一电荷的狄拉克量子化条件。至于非平凡规范变换下的规范不变性产生的物理参量的量子化的更多例子，可在贾基夫的论文(1985)中找到。

从上面的讨论我们知道，具有简单的非阿贝尔规范群(其中变换是非平凡的)的规范不变理论总含有量子化条件。如果对称群被标量场自发破缺到 U(1)群(其变换等同于电磁规范变换)，那么这些量子化条件都是平凡的吗？特霍夫特(1974)和波利亚科夫(A. Polyakov, 1974)发现，在 SU(2)规范群的情形中，这些量子化条件将会保留，因为光滑 SU(2)单极子将会代替奇异的狄拉克单极子，在原点可能是匀称的构形(configuration)，远离原点后就可能成为单极子。后来，人们找到光滑单极子的具有有限经典能的静态解，并且相信相对于扰动解是稳定的。物理学家试图把光滑单极子和三维空间中的非线性场论的孤立子等同起来，并用半经典方法描述它的量子态。[53]

特霍夫特—波利亚科夫单极子及其推广，必定在每一个大统一理论中都存在，因为它们都建立在简单对称群的基础之上，后者会因 U(1)因子的存在而自发破缺。因此，单极子的存在的实验证据的缺乏和成问题的理论推论(比如质子衰变)，对模型的建立有很强的制约作用。

另一个非阿贝尔规范理论回应有限规范变换的有趣推论，涉及真空结构的新概念。

在纯粹非阿贝尔规范理论中，人们固定了规范之后(比如采用规范 $A_0 = 0$)，另外还有一些剩余规范自由度，比如那些有限规范变换 $U(r)$，它们只跟空间变量有关。因此，由 $F_{\mu\nu} = 0$ 定义的真空，用 $\lim U_{(r\to\infty)} = U_\infty$［这里的 U_∞ 是指整体(与位置无关)规范变换］，把规范势限制到“纯规范”和零能组态 $A_i(r) = -U^{-1}(r)\partial_i U(r)$。由于 Π_3(紧致非阿贝尔规范群) $= Z$（见附录 2)，对非阿贝尔规范理论来说，必定存在一个在拓扑学上可区分的但在能量上简并的无穷多态真空，它能用 $A^{(n)} = U_n^{-1}\nabla U_n$ 表示，n 从 $-\infty$ 一直到 $+\infty$。因此物理真空的量子态应该是波函数 $\Psi_n(A)$ 的叠加，即 $\Psi(A) = \sum_n e^{in\theta}\Psi_n(A)$，其中每一个 $\Psi_n(A)$ 都位于经典零能组态 $A^{(n)}$ 附近。它们相对于同伦平凡变换一定是规范不变的。但是非平凡变换平移 n：$S_n\Psi_{n'}' = \Psi_{n+n'}$（这里的 S_n 是幺正算符)，意味着有限规范变换 U_n 属于第 n 同伦类。因为 $S_n\Psi(A) = e^{-in\theta}\Psi(A)$，所以物理真空由 θ 刻画，θ 即著名的真空角，这种真空就叫 θ 真空。[54]

[310]

纯规范势 $A^{(n)} = U_n^{-1}\nabla U_n$ 类比于在具有周期势的量子力学问题中无穷简并的零能量组态，而在拓扑真空态对有限规范变换的反应中相位的出现，类比于周期势问题中波函数由此得到一个相位的移动。和在量子力学问题中一样，如果不同拓扑真空之间彼此存在隧穿，其简并度就会增加。在纯粹非阿贝尔规范理论中，这种隧穿归因于瞬子(instantons)。[55]当无质量费米子包括在理论中时，隧穿被抑制，真空保持简并，而 θ 角没有物理意义。[56]在非阿贝尔规范理论中，比如量子色动力学［其中费米子并非无质量(见 11.1 节)］中，θ 角在理论上应当是可观测的，虽然在实验上它还没有被观测到。[57]

总之，依靠在同伦上是非平凡的大的或有限的变换的规范不变性

概念，开启了规范理论的一个新视角，有助于揭示或反思某些在本质上是整体的非阿贝尔规范理论的新特点，并因此提供了把用场论描述的物理世界的局部观和整体观结合起来的强有力手段。

10.5 未解决的问题

虽然规范理论的牢固基础在近四十年已经具备，但是许多概念问题仍未成定论，表明这些基础并非坚如磐石，也不是没有毛病的。最重要的概念问题有三个。

Ⅰ. 戈德斯通—希格斯标量系统的本体论地位

[311]

正如我们已经看到的，就涉及的基本思想而论，戈德斯通关于自发对称性破缺的工作和南部的工作没有什么不同。不过，戈德斯通采用标量玻色子作为基本粒子，并在玻色子系统中探索自发对称性破缺的条件和结果；而在南部的框架里，标量玻色子是导出来的，因为这些是复合模式，所以仅仅表现为费米子系统中对称性破缺的结果。戈德斯通模型的一个优点是可重正化性。这使其在找到非线性系统的反对称解的存在条件方面容易得多，然而，比这更有趣的是通过在自发对称性破缺研究中引入基本标量系统所产生的新特点。

首先，戈德斯通玻色子系统中的对称性破缺的一个征兆，是有质量标量粒子的不完备多重态的出现，即所谓的希格斯玻色子(Higgs bosons)。在南部的框架中，没有明确的对称性破缺，无质量的无自旋玻色子也是可能的。因此与南部的方法相比，戈德斯通的方法引出了多余的理论结构，即有质量标量玻色子。说它是多余的是因为它不为费米子系统恢复对称性所必需。

第二，在研究对称性破缺时，基本标量系统的引入给对称性破缺

强加了一个双重结构。也就是说，费米子系统的自发对称性破缺没有被其自身的非线性动力学结构确定下来。说得更确切些，费米子系统的自发对称性破缺通过汤川耦合和规范耦合被标量玻色子主系统中的对称性破缺所诱导。对称性破缺的这种双重结构导致了标准模型的特殊性(见 11.1 节)。那就是，除了理论上固有的动力学部分(用于解释和预言观测)之外，还有一个可任意调节部分(用于实际物理状态的现象学上的选择)。

戈德斯通—希格斯标量系统的引入，已经开启了我们理解物理世界——从基本粒子谱系及其相互作用到整个宇宙的结构和演化——新的可能性。特别是，希格斯场与引起规范理论纲领的两个最革命的方面密切相关。首先，它首次提出质量(不仅是规范量子的质量，还有费米子的质量)是如何通过相互作用产生的：规范相互作用产生规范量子的质量，希格斯场和费米子场之间的所谓汤川相互作用产生费米子的质量。汤川相互作用是规范不变的，但不能从规范原理中导[312] 出。也就是说，虽然汤川相互作用未与规范原理相矛盾，但它自身已越出规范理论的框架。事实上，理解汤川相互作用的本质，要求在场论研究方面有一个新的方向。其次，额外的希格斯场结构在一个新类型的真空中起了作用，这有点类似于老式以太，其作用像连续背景介质，并弥漫于整个时空之中。在某些模型中，有人提出这种结构是宇宙学常量和宇宙相变的原因。因而问题在于这个以太来自哪里或希格斯场的起源是什么。

因此为了成功地贯彻规范理论纲领，彻底详细地理解戈德斯通—希格斯标量系统的本质及其动力学，似乎比理解规范场论本身更重要，正如理解规范场对贯彻量子场纲领来说比理解量子场论本身更重要。于是，标量系统的本体论地位这个严肃问题就出

现了。

注意，在标量系统和超导系统之间存在很大差异。在超导系统中，反对称相位、对称相位和这两种相位之间的相变都是真实的；然而，在标量系统中，戈德斯通玻色子是非物理的，我们观测不到希格斯玻色子，而对称解几乎没有引起物理学家的注意。非物理戈德斯通标量和其他非物理学标量、协变鬼和法捷耶夫—波波夫鬼，都深深地卷入非阿贝尔规范理论的理论结构即规范玻色子的描述中。关于这些非物理自由度的一个不可避免的物理问题涉及它们与自然的关系：它们代表了物理实在吗？或者仅仅是没有直接的物理意义的用于信息编码的辅助结构？

工具主义者把它作为建立诸如标准模型之类的模型的特殊设计，以便所要求的像 W 粒子和中性流之类的东西都能观测到。他们并没有严肃地考虑包括希格斯玻色子这些特殊设计的所有含义。但是工具主义者不得不面对接下来的问题。在这些结构中编码起来的信息的地位如何？我们有可能直接得到相关信息而不求助于虚拟设计吗？也就是说，我们能否重新建构一个自洽的理论，其综合力和解释力、预测力均等同于标准模型，又没有这种深深植根在标准模型中的非物理自由度？

如果我们采取实在论的立场，那么寻找希格斯粒子，寻找标量系统的对称解，以及寻找促使系统从其对称相位到反对称相位的动因，[313] 都将是严肃的物理学问题。再者，还有关于标量粒子的本质的另一个问题：它们是基本的还是复合的？

有些物理学家感到只有在唯象模型中才能把标量粒子看作基本的，而在基础理论中它们应当从费米子导出。对他们来说，戈德斯通的方法是对南部更有雄心的纲领的一种再处理，而且包括人工色

(technicolor)的思想在内的动力学对称性破缺思想，似乎比戈德斯通的方法更有吸引力。不过，标量系统在标准模型中所起的主要作用似乎支持另一种观点，这种观点在20世纪50年代由西岛和彦(1957，1958)、齐默尔曼(1958)和哈格(1958)，以及在20世纪60年代早中期由温伯格(1965a)和其他一些人广泛探讨过，也就是认为就散射理论而言，基本粒子和复合粒子之间的差异不存在。

II．没有新的物理学了吗？

习惯上，物理学家相信量子场论会突破短程限制，而随着越来越强大的实验仪器探索越来越高的能区，新物理学会出现(见8.3节)。[58]然而，随着渐近自由的发现，一些物理学家争辩说，由于在渐近自由理论中裸耦合是有限的并且逐渐消散，那么总体上没有无穷大了，并且这重申了四维量子场论的连贯性。尤其是，高能有效耦合的减少意味着在短程范围没有新物理学会兴起(Gross，1992)。

所有这些说法都很吸引人。不过，问题在于四维中的渐近自由规范理论的存在性没有严格证明。这类证明必须建立在重正化群的论证基础之上，并要求对紫外稳定不动点的存在进行证明。然而在数学上还没有完成不动点存在的证明。这个困难也表现在其以一种更令人难以忍受的方式进入诸如红外动力学或低能物理学这些较不重要的方面：不仅没有四维时空中色禁闭的精确解，并且通过渐近自由理论对低能 π 介子—核子相互作用的解释，看来也几乎做不到。[59]

III．可重正化性被证明了吗？

[314] 有质量规范玻色子的规范理论的可重正化性的所有证明，至少涉及一个额外粒子，即希格斯粒子。因此其可重正化性证明中一个必不可少的部分，应该是对希格斯质量有一个说法。然而直到现在，还完全没有这种说明，我们也没有来自实验的有关它的定量方面的任

何线索。更糟的是，物理学家对希格斯粒子的存在或找到它的可能性，也存在意见分歧。他们中的许多人感到世界比我们用希格斯粒子概念所理解的更复杂，虽然它已能很好地对目前实验的情形参量化(Veltman，1992)。因此，在更完满的意义上，有质量规范理论的可重正化性的证明仍然处于可疑阶段。

在更深的层次上，量子场论的可重正化性的严格证明，只能通过重正化群的方法得到。如果一个理论有一个重正化群的不动点，那么这个理论是可重正化的，否则就不是。不过，与渐近自由的证明的情况一样糟，不动点存在的证明在数学上从未实现。因此声称非阿贝尔规范理论的可重正化性已被证明，有些言过其实。而且，近来的发展已经提出了量子场论中不可重正化相互作用的合法性问题(见 11.4 节)，因此，对量子场论的连贯性的理解，以及在构建一个连贯的量子场论中可重正化性证明的作用，提出了一系列严峻问题。

注释

1. 这个术语是贝克和格拉肖(1962)发明的，因为他们注意到对称性破缺的机制并不需要在拉格朗日函数中有任何明确的质量项以明显破坏规范不变性。关于这个问题的更详细的历史考察可参阅 Brown and Cao(1991)。

2. 维格纳强调物理状态和物理学定律之间差异的重要性，并且断言对称性原理“仅适用于我们关于自然界知识的次级概念，即所谓的自然律”(Houtappel, Van Dam, and Wigner, 1965；另见维格纳更早的论文，1949，1964a，b)。维格纳的评论夸大了这种差异，因为物理状态(比如，基本粒子的状态)可以构成对称群表象。事实上，维格纳的著作(1931)涉及把群论应用到原子态。采用维格纳关于对称性原理的可应用性的严格说法，将会阻断通向自发对称性破缺的推理线索。正如我们下文将看到的，自发对称性破缺的概念完全建立在描述物理系统的方程的对称性和系统的物理状态的更严格的对称性的比较之上。

3. 能在 L. A. Radicati(1987)中找到它们。

4. 在铁磁体的情形中，这个参量也是一个运动常量，是自发对称性破缺的非典型简化特性。在反铁磁情况下，首次被安德森(1952)处理得更有代表性：宏观参量，虽然不是运动常量，但是和自旋群的生成元仍有联系，并且表明了对称性破缺的程度。 [315]

5. 引自 Brown and Rechenberg(1988)。

6. 他这么做并非偶然，因为早在 1922 年还是慕尼黑大学一年级学生的时候，他就进行了湍流问题的研究。后来，在其老师索末菲的建议下，海森伯把湍流的开端问题(自发对称性破缺的一个例子)作为毕业论文。见 Cassidy and Rechenberg(1985)。

7. 真空具有很大的同位旋和其他内禀量子数，或者用温策尔(1956)的术语来说，真空有大量各种各样的"乱真子"(spurions)，因此内禀量子数能从真空中根据需要抽取出来。这个思想即使对温策尔而言似乎也是不自然的，并且受到泡利的严厉批评。在日内瓦会议上，盖尔曼写信给海森伯，重述泡利的如下反对意见："取决于你想描述多少个粒子，N 个，那么你就要 2^N 个真空……我假定你能这样做，但它似乎很复杂并且很像是要增加另外一个场。"见海森伯的评论之后盖尔曼的讨论(1958)。

8. Dürr, Heisenberg, Mitter, Schlieder, and Yamazaki(1959).

9. 在 1960 年的罗切斯特会议上，海森伯发展了这些观点并重申用"世界态"取代"真空"(1960)。

10. 在关于非线性场论的论文(1961)中，海森伯从自己早期关于真空的简并性的更激进的立场退出，并开始诉诸非简并真空的大范围涨落的思想。因为这个原因，并且也由于他不能把他自己的工作和南部以及戈德斯通的工作联系起来，或把它与杨—米尔斯理论联系起来，海森伯对自发对称性破缺的发展的影响下降了。在他雄心勃勃的纯理论的非线性场论毫无结果之后，海森伯偏离了粒子物理学的主流。及至此时，他对自发对称性破缺的发展的主要贡献看上去几乎被完全遗忘。

11. 这场争论实际上可以追溯到 1956 年，BCS 理论出现的前一年。当时白金汉(M. J. Buckingham, 1957)向巴丁挑战，后者在 1955 年就提出，假定光跃迁的矩阵元跟它们没有能隙时完全相同，能隙模型可能产生迈斯纳效应。白金汉证明了它们只有在朗道规范中才是相同的，而在一般情况中不同。沙弗罗特(Max. R. Schafroth)在 1958 年反驳了迈斯纳效应的 BCS 推导。他指出，早在 1951 年他就证明了。涉及违背规范不变性理论中的任何小差错一般都会产生伪迈斯纳效应，一旦规范不变性得到恢复，伪迈斯纳效应也就消失(Schafroth, 1951)。

12. Bardeen(1957), Anderson(1958a, b), Pines and Schrieffer(1958), Wentzel(1958, 1959), Rickaysen(1958), Blatt, Matsubara, and May(1959).

13. Bogoliubov(1958), Nambu(1960).

14. 应当注意到超导相只是完备系统的一部分，其总体描述一定是规范不变的。

15. γ_5 是 4 个狄拉克矩阵 γ_μ 的积，并且和它们中的任何一个都是反对易的。因此它和狄拉克方程 $D\psi = m\psi$ 中狄拉克算符 D 也是反对易的，因为 D 在 γ_μ 中是线性的，但是它和质量 m 是对易的。因此只有无质量狄拉克方程是 γ_5 不变的。

16. Nambu and Jona-Lasinio(1961a, b).

17. 在其中西部会议的论文中，他说道："在金属中的费米电子海类比于真空中的狄拉克电子海，在这两种情况下我们说的都是电子和空穴。"(1960c)

18. 南部说得很清楚："我们的 γ_5 不变性理论能够近似于用赝标量介子所作的现象学描述……其原因在于真空的简并性和建立于其上的世界。"(Nambu and Jona-Lasinio, 1961a)

19. 参阅 Nambu(1989)。

20. 在超导性中，库仑相互作用促使集体激发进入有质量等离子体，南部本可应用类似的推理得到 π 介子质量。但是这种严格对称的自发破缺，虽然允许 π 介子有非零质量，但这意味着轴矢流的守恒，这反过来意味着 π 介子衰变率趋于零，这与观测不符。受此物理考虑限制，南部放弃了这一类比并接受了部分守恒轴矢流思想，后者是盖尔曼和利维(Maurice Levy)等人提出来的(见 8.6 节)。因此，南部采纳了在小裸核子质量形式中 γ_5 不变性被明确破缺的方式(Nambu and Jona-Lasinio, 1961a)。[316]

21. 盖尔曼声称他和合作者在他们关于其发散是由低质量 π 介子极点所支配的部分守恒轴矢流的工作中，独立地找到了南部—戈德斯通玻色子。他在 1987 还宣布："按照部分守恒轴矢流，当[轴]流的发散趋于零时，m_π 比 m_n 更快趋于零。在极限情况下，存在'南部—戈德斯通'机制的实现；这些作者在差不多相同时候独立地发展了这一机制。"在 1989 年他再次声称："在严格守恒的极限情况下，π 介子将没有质量，并且这是南部—戈德斯通机制的实现。"

部分守恒轴矢流的极限情况和南部—戈德斯通机制的冲突多少有些误导，因为产生这两者的物理学思想不同。南部—戈德斯通机制是建立在简并真空态思想基础之上的，这

种真空态是非线性动力学系统的稳定的非对称解。因此，对称性在方程解的层次上而不是在动力学定律的层次上破缺。在部分守恒轴矢流的情形中，既没有非线性也没有真空的简并性作为其描述特点，而且对称性是在动力学方程的层次上破缺。一个说明性事实是，在部分守恒轴矢流的框架中，一旦对称性破缺，就没有无质量无自旋玻色子。一旦获得无质量玻色子(通过取部分守恒轴矢流的守恒极限)，就完全没有对称性破缺。和这种情况形成鲜明对比的是，在南部—戈德斯通机制中无质量标量玻色子的出现是作为对称性破缺的结果。这些无质量玻色子和反对称解是共存的，以使整个系统的对称性得以恢复。总之，南部—戈德斯通玻色子是具有连续对称性群的理论中自发对称性破缺的结果，而在盖尔曼及其合作者的工作中的无质量 π 介子和自发对称性破缺没有关系。考察盖尔曼和利维(1960)以及伯恩斯坦、富比尼、盖尔曼和蒂林的论文(J. Bernstein, S. Fubini, Gell-Mann, and W. Thirring, 1960)，我在这些论文中没有发现关于真空简并性、反对称解的存在或通过无质量场恢复对称性的说法。他们关于有质量的或可能无质量的 π 介子的区分发生在完全不同的语境中，与南部—戈德斯通玻色子无关。

22. 在寻找量子场论的"超导"解的南部方法和戈德斯通方法之间存在一个重大差异。南部的方法是"微观的"，与 BCS—博戈留波夫理论相差不多。戈德斯通的方法是"宏观的"，在"现象学的"意义上与朗道—金兹堡理论(Landau - Ginzburg theory)相差不多。在戈德斯通模型中，玻色子直接来自集体激发，正如能隙参数在朗道—金兹堡理论中的作用一样，它们不是从微观层次的原场的相互作用中导出的。这种通向自发对称性破缺的宏观方法在标准模型中也是标准方法，并且有助于解释它的特殊性质：模型的一部分给出动力学，而另一部分自发对称性破缺对实际物理态作出现象学的选择。第一部分在理论上是固定了的，预言可观测结果；第二部分是可任意调节的。在某些方面，戈德斯通的工作避开了南部更为雄心勃勃的纲领。不过，从我们对自然界一些基本成分和它们的相互作用的无知来看，这种再处理在战略上是有价值的。

23. 盖尔曼声称希格斯机制也是对软质量问题的一个解答(1987, 1989)。软质量的思想是对可重正化性和对称性破缺(包括得到近似的整体对称性)这样一些当代主要成见的回应。不过软质量机制的最初公式只是简单地把规范玻色子质量项增加到规范不变的格拉朗日函数中，这一项既破坏了规范不变性又破坏了可重正化性。只是在后来的发展中，才在自发对称性破缺的思想和施温格的强耦合的思想基础之上找到了一种机制，通过这一机制规范玻色子获得了质量而没有破坏拉格朗日函数的对称性。不过，规范玻色子通过这个机制获得的质量绝不是软的，相反，它们极其硬。 [317]

24. 正如我在前面提到的，超导性的朗道—金兹堡理论(1950)已经明确使用自发对称性破缺，并且包含可被进一步发展并整合进规范理论框架的自发对称性破缺特点。比如，复标量场 ψ 的相模(在 BCS 理论中其意义被澄清为"能隙"的序参量)被朗道和金兹堡引进来描述超导体中的电子合作态，被证明是无质量南部—戈德斯通模的原型，而且振幅模变成希格斯模的原型。不过，因为朗道—金兹堡理论是一个宏观理论，其相关性没有引起粒子物理学家的注意。描述超导性和量子场论的理论结构之间的类似，是在有关超导性的微观理论即 BCS 理论于 1957 年发表后，才变得明显的。

25. 这是物理学概念基础中那些轮回的辩证迁移之一。我们遵循它，在经典层次上，从笛卡儿到牛顿以及从麦克斯韦到爱因斯坦；在量子层次上，从海森伯和薛定谔到狄拉克。狄拉克用负能电子海和其他费米子海填满真空，哥本哈根学派的成员及其追随者尽力从负能海中解放出真空(见 Weisskopf, 1983)。从 19 世纪 40 年代后期起，认为真空是一种空虚变得流行起来，直到简并真空是充盈的这种观点与自发对称性破缺联系起来并再次成为主流。

26. Bloom *et al*. (1969).

27. 为了轻松理解如何从部分子模型导出比约肯标度无关性，有关技术问题的说明，见 Pais(1986)。

28. 卡伦(1970)的标题是"标量场论中的比约肯标度不变性"，西曼齐克(1970)的标题是"场论中小范围行为和方次计数"。

29. Symanzik(1971)和 Callan(1972)。

30. 见 Jackiw, Van Royen and West(1970), Leutwyler and Stern(1970), Frishman

(1971), and Gross and Treiman(1971)。

31. Fritzsch, Gell-Mann, and Leutwyler(1973).

32. 这是由波利策(1973)以及格罗斯和学生维尔切克(1973a, b)计算出来的。

33. 相同类型的贡献由 $W^{\pm}$ 和 Z 玻色子提供，这是通过施温格—安德森—希格斯机制产生并传递弱荷。但是这种贡献与无质量胶子的贡献相比，被这些有质量玻色子的巨大质量极大地抑制住了。另一方面，格罗斯和维尔切克(1973a)起先担心自旋为零的希格斯玻色子会引起附加屏蔽，并损坏非阿贝尔规范理论自发破缺中渐近自由的可能性。但是如果希格斯玻色子质量很大，那么同样的抑制机制会使其贡献不可忽略。

34. 例如，见 Wilson(1974)和 Creuz(1981)。

35. 见 Umezawa and Kawabe(1949a, b), Feldman(1949), Neuman and Furry (1949), Case(1949), Kinoshita(1950), Umezawa and Kamefuchi(1951), and Sakata, Umezawa, and Kamefuchi(1952)。

36. 见第 9 章的注释 4。

37. 这个出版物(1963)基于他在关于引力的相对论会议上的演讲录音带，此会议于 1962 年 7 月在波兰的亚布翁纳举行。

38. 由于费恩曼把幺正性作为圈和树之间的联络，所以他未能超出圈的范围去讨论这个问题。

39. 矢量传播子的这个形式是在所谓幺正规范中得到的。正如长期以来认识到的，用传播子的这一形式，包含 $q^{\mu}q_{\nu}/\mu^2$ 的项对 S 矩阵起了不可重正化的作用。

40. 因此，鬼传播子有一个方位或方向，这是费恩曼没有提到的。韦尔特曼后来评论 [318] 了德威特对规范理论中鬼概念的贡献说道："这个鬼现在叫做法捷耶夫—波波夫鬼，多少有些不合理。"(1973)

41. 正如我们在前面所注意到的，这是从幺正规范到朗道规范的改变。

42. 在自由场方法和有质量量子电动力学中斯蒂克尔伯格公式(1938)之间，在两者都引入附加标量场的意义上存在某些明显的相似之处。不过两者的区别也是本质的。首先，在有质量量子电动力学中斯蒂克尔伯格公式的纵向分量是退耦的，但是从贝尔—特赖曼变换导出的新费恩曼规则包括新顶点和鬼场。第二，斯蒂克尔伯格标量场在规范变换下以某种方法进行变换，而韦尔特曼标量场自身是规范变化的。第三，由于规范不变性被引入的代表电磁场的质量项所破坏，而被标量场部分恢复，所以斯蒂克尔伯格方法会导致希格斯机制的发现，而不是法捷耶夫—波波夫鬼，后者是由规范不变性的破缺来补偿的。我感激韦尔特曼教授，因为他帮我澄清了这个复杂问题。

43. Veltman(1969)以及 Reiff and Veltman(1969)。

44. 后来，斯拉夫诺夫(1972)和泰勒(1971)通过使用路径积分公式，为无质量杨—米尔斯理论导出了广义沃德恒等式，这就是著名的斯拉夫诺夫—泰勒恒等式。这些恒等式和有质量理论的韦尔特曼恒等式相似，并且也能被规范对称性自发破缺的理论记录下来。它们包含理论的所有综合内容，因为从它们那里可推导出拉格朗日函数的对称性。

45. 这些明确的鬼费恩曼规则是相对于法捷耶夫和波波夫的，也是相对于弗拉德金和秋京的，后者按确定的方法写下拉格朗日函数的鬼部分。

46. Lee(1969), Gervais and Lee(1969), 以及 Symanzik(1969, 1970)。

47. Lee(1972).

48. 维数正规化遵守大多数对称性，但涉及 γ_5 矩阵的标度对称性和手征对称性除外，因为这两种对称性是维度确定的(γ_5 矩阵只能在四维时空中定义)。

49. 事实上，规范不变性只是场论的可重正化性的两个关键因素之一。另一个是耦合常数的维度。弱相互作用中的希格斯型规范理论可重正化之所以是可能的，就因为耦合常数 G_w 的维度(M^{-2})在直接的费米耦合中能转移到有质量中间介子($\sim q^{\mu}q_{\nu}/M^2q^2$)的传播子。不过，这种转移不可应用于有关引力的量子理论。在引力的规范理论中，因为引力的长程特性，传递引力相互作用的规范量子(引力子)是无质量的。因此，尽管有韦尔特曼和特霍夫特的工作，引力量子理论的可重正化性仍然是对理论物理学家的严峻挑战。

50. 除了卡布雷拉(B. Cabrera, 1982)观测到单极子的独一无二的报告外。

51. 杨振宁(1970)论证说："[在]作用于电荷 e_j 的电场 ψ_j 上的时空无关的规范变换中：

$$\psi_j \rightarrow \psi_j' = \psi_j \exp(ie\alpha) \tag{1}$$

……如果不同场的不同 $e_j(=e_1, e_2, \cdots)$ 是彼此不可通约的，那么变换(1)对于 α 的所有实值都不同，并且必须定义规范群以包括 α 的所有实值。因此，群不紧致。

另一方面，如果所有不同的 e_j 是 e 的整数倍，e 是电荷的普适单位，那么乘以 $2\pi/e$ 的整数倍，α 的两个值还是不同，变换(1)对任何场 ψ_j 都是相同的。换言之，如果 α 是相同的模 $2\pi/e$，那么两个变换(1)是不可分辨的。因此由(1)定义的规范群是紧致的。"

52. 见 Wu and Yang(1975)。

53. 见 Jackiw and Rebbi(1976)以及 Colman(1977)。

54. 见 Callan, Dashen, and Gross(1976)，以及 Jackiw and Rebbi(1976)。

55. Belavin, Polyakov, Schwartz, and Tyupkin(1975).

56. 见 't Hooft(1976a, b)，Jackiw and Rebbi(1976)。 [319]

57. 关于中子的偶极子动量严格的实验极限的分析表明 $\theta \leqslant 10^{-9}$。见 Crewther, Divecchia, Veneziano and Witten(1979)。

58. 诸如 Schwinger(1948b, 1970)和 Gell-Mann and Low(1954)。

59. 虽然在把这些困难理解为量子色动力学的结果这个方向上有些进步，但这些进步是通过使用各种有效作用方法在格点规范理论的框架内得到的。见 Wilson(1974)以及 Kogut and Susskind(1975)。

第11章
规范场纲领

[320] 正如当代一些有影响力的物理学家所评论的，1971年问世的规范不变有质量矢量介子理论的可重正化性的乌得勒支证明(Utrecht proof)，“将以最深刻的方式改变我们关于规范场论的思维方式”(Lee，1972)，它“引起强烈的轰动，使统一问题成为集中研究的主题”(Pais，1986)。更准确地说，怀着对前景巨大改变的兴奋之情，粒子物理学共同体的自信心很快就在其集体意识中建立起来。并且相信由规范原理所确定的量子场体系，是一个用统一方式描述基本相互作用的自洽的、强有力的概念框架。新视野下的一项最新成果是所谓的标准模型(standard model)的兴起，它由弱电理论和量子色动力学(QCD)组成(11.1节)。该模型在实验上所取得的成功及其强有力的基本概念鼓舞了物理学家，他们假定规范原理具有普适性，并且为把模型扩大到大统一理论和引力上付出了很多努力(11.2节)。

规范场纲领(GFP)的出现引出了对20世纪场论发展的辩证理解：把规范场纲领作为几何纲领和量子场纲领的综合。11.3节给出

这样一种理解的某些辩护。但是场论的辩证发展并没有终止于黑格尔式的闭合终端(一个最终理论或一个封闭的理论框架)。11.4 节中给出了一些关于规范场纲领之所以停滞的讨论和关于场论研究新方向的讨论。

11.1 标准模型的兴起

"标准模型"这个术语是派斯和特赖曼在 1975 年首次创造的，用来指称具有 4 种夸克的弱电理论。在其后来的使用中，它指强子和轻子结合在一起的 6 夸克绘景，及其用弱电理论和量子色动力学描述的动力学。从历史的观点看，标准模型的兴起是如下三方面发展的结果：(i)概念框架的建立，(ii)其协调性的证明，以及(iii)模型的建立。标准模型建立于其中的一般概念框架是把规范耦合思想和对称性破缺思想相结合的强子的夸克模型。夸克模型最初是由盖尔曼(1964a)和茨威格(George Zweig, 1964a, b)提出，他们用三种(后来称为味)夸克形成 SU(3)群的基底，作为使强子中破缺的 SU(3)对称性结合起来的一种方法，让夸克(具有重子数为 1/3、自旋为 1/2 以及分数电荷的强子的一部分)和轻子一起，被当作自然界的微观结构的基本成分。然而夸克的实在性是在几年后才确立的，要不然完全没有一个人(甚至盖尔曼自己)能严格地把夸克模型作为对物理世界理论化的本体论基础。建立夸克实在性的关键一步是在 1969 年取得的，当时在斯坦福直线加速器中心进行深度非弹性散射实验以探测强子的短程结构。所观测到的比约肯标度无关性表明，强子由自由点状成分或部分子组成(见 10.2 节)；当有些实验数据按照流的算符乘积分析时，其中有些部分子表现出是带电的，具有 1/3 重子数、1/2 自旋。也就是说，它们看起来像夸克。进一步的数据也表明，它们

[321]

跟赋予夸克的电荷量相一致。[1]这些初步的但很关键的进展，自然使一些先锋物理学家确信夸克的实在性，并鼓励他们在对亚原子世界进行概念化时使用夸克模型。

规范耦合概念是由杨振宁和米尔斯提出的，并进一步由内山、施温格、格拉肖等等许多人所发展(见9.1节)。与这个规范对称性概念互补的是各种对称性破缺的概念：由海森伯、南部、戈德斯通、安德森等等许多人发展起来的自发破缺概念(10.1节)，以及由阿德勒、贝尔与贾基夫等等许多人发展的反常破缺的概念(8.7节)。这些概念在建立框架的协调性和模型建构上起到了不可或缺的作用。但这个概念框架最重要和最深刻的部分，是标度不变性的反常破缺的一个结果，这种标度不变性体现在重正化群的概念之中(见8.8节)。重正化群的思想假设了在不同能量标度上物理相互作用的结构之间因果联结的具体类型，否则就没有巡行耦合的思想，从而也就没有渐近[322]自由和大统一的思想，可重正化性的严格证明也将不可能。

非阿贝尔规范理论的协调性证明

至于有质量矢量理论，安德森—希格斯机制的引入，即使不是决定性的一步，也是朝正确方向迈出的重要一步。之所以重要是因为，运用安德森—希格斯机制能使规范玻色子获得质量(这对拯救弱相互作用领域的现象起关键作用)，从而与精确的而不是近似的或部分的动力学对称性相容。这使可重正化性的证明变得容易得多。无论如何，认识到这个优点，加上人们普遍都相信具有未破缺对称性的杨—米尔斯理论是可重正化的，最终产生了一种天真的信念，认为希格斯型有质量杨—米尔斯理论是先验可重正化的。

说此信念是天真的是因为其说服力仅仅建立在简单方次计数论证基础上，因此证据不足。这里我们不要忘记，规范不变性和可重正

化性之间的关系问题，比 20 世纪 50 年代和 60 年代杨振宁和米尔斯、格拉肖、科马尔和萨拉姆、温伯格等等许多人所想的要尖锐得多（见 10.3 节）。对于需要可重正化性的理论，规范不变性既非充分条件（比如引力的情形），也非必要条件（比如中性矢量介子理论的情形），虽然它肯定对模型的建立设置了严格限制。因此，非阿贝尔规范理论的可重正化性证明对理论学家来说是一个巨大的挑战。它要求理论物理学家对其作严肃和艰难的探索，这正是从李政道和杨振宁开始，经过费恩曼、德威特、法捷耶夫与波波夫，直到韦尔特曼和特霍夫特所做的工作。他们按照协调性方法引入了已有物理学原理所要求的复杂的非物理的自由度系统，对这个理论进行了量子化。他们推导出费恩曼规则和沃德恒等式，发明了可重正化规范，并证明了幺正性。最后，他们发明了规范不变的规则化方案。没有这些研究，非阿贝尔规范理论的可重正化性证明是不可能的，并且那些建立在对称性论证和简单方次计数论证上的所有先验式的信念和推测，都是无根据和空洞的。

模型的建立：弱电理论

在 9.3 节中可以看到，最早在非阿贝尔规范理论的框架内，物理学家为建立描述基本相互作用的模型作了种种努力。就弱相互作用而言，这些努力中有两项主要成就似乎是永久性的，它们是通过经验数据确定的（有质量玻色子所要求的短程行为，手征不对称性所要求的宇称不守恒），即：(i)把 $SU(2)\times U(1)$ 作为统一描述弱相互作用和电磁相互作用的规范群，(ii)采用安德森—希格斯机制产生质量。 [323]

然而，这些努力在理论上是十分脆弱的，没有给出模型的协调一致的量子化。如果没有适当的费恩曼规则的指示，那么关于可重正化性的思考，或它是否会被自发对称性破缺所破坏，在概念上都是空

的。因为这个原因，正如韦尔特曼在其有影响力的评论(1973)中所谈到的，这些努力“是不能令人信服的，而这条研究路线直到最近才产生进一步的成果”。韦尔特曼的评论没有偏离科尔曼(1979)，沙利文、凯斯特、怀特和克恩(D. Sullivan, D. Koester, D. H. White and K. Kern, 1980)以及派斯所支持的说法。派斯注意到温伯格提出的模型“在 1967 年到 1969 年间的文献中很少有人引用。在维也纳(1968 年)和基辅(1970 年)的两年一次的罗切斯特会议上，大会报告起草人的谈话(常常专门讨论当时流行什么)，甚至也没有提到 SU(2) × U(1)”(Pais, 1986)。[2]

正如派斯注意到的，当可重正化性的乌得勒支证明于 1971 年 6 月在由韦尔特曼组织的阿姆斯特丹会议上提出后，巨大的变化开始了。除了有质量杨—米尔斯理论的可重正化性证明之外，特霍夫特(1971b, c)还提出了三个模型。特霍夫特的模型之一，是和先前由温伯格和萨拉姆所提出而又被遗忘的模型等同的，而温伯格和萨拉姆的模型除了希格斯机制外，和格拉肖(1961)提出的模型是等效的。

这个(后来成为)著名模型的核心是其规范群，该群是格拉肖首次提出来的。由于在唯象理论上荷流和弱相互作用中宇称不守恒的限制，人们认为弱同位旋群 $SU(2)_T$ 会引起弱耦合。关于弱耦合，左手轻子场被指派为双重态(doublets)

$$L_{e(\mu)} = [(1-\gamma_5)/2]\begin{vmatrix} \nu_{e(\mu)} \\ e(\mu) \end{vmatrix},$$

右手轻子场被指派为单态(singlets)($R_{e(\mu)} = [(1+\gamma_5)/2]e(\mu)$)，而规范场被指派为三重态(triplets, $W_\mu^\pm$ 和 W_μ^0)。为了统一弱相互作用和电磁相互作用，人们引入了一个新的规范群——“弱超荷”群

$U(1)_Y$——和相关的矢量场 B_μ，其相应的量子数满足关系式 $Q = T_3 + Y/2$，其中 Q 是电荷，T_3 是弱同位旋 T 的第三分量，并且 T 和 Y 相互对易。按这种方式，联合相互作用可以用 $SU(2)_T \times U(1)_Y$ [324] 乘积给定的新规范群的规范玻色子描述。在不破坏对称性的前提下，此规范玻色子的质量可以通过安德森—希格斯机制获得。

可以探索性地把 $Q = T_3 + Y/2$ 理解成这样一种陈述，即光子场 A_μ 是 W^0_μ 和 B_μ 的某种线性组合。那么正交组合 Z_μ 就代表自旋为 1 的中性有质量粒子：

$$A_\mu = \sin\theta_W W^0_\mu + \cos\theta_W B_\mu$$

和

$$Z_\mu = \cos\theta_W W^0_\mu - \sin\theta_W B_\mu$$

其中 θ_W 是温伯格角（$e = g\sin\theta_W = g'\cos\theta_W$，$g$ 和 g' 是 W_μ 和 B_μ 各自的耦合常数）。因此，把规范群作为没有关联的群的乘积，这样的结合是通过规范场的混合实现的。

建立模型的下一步是由反常论证决定的。在有手征流的特霍夫特模型中，不可避免的手征反常肯定要破坏其可重正化性，这是在 1972 年 1 月的奥尔赛会议上由巴丁首次指出的，当时特霍夫特认为反常没有危害。[3]受到奥尔赛会议上争论的激发，布夏、伊奥普洛和梅耶尔博士（C. Bouchiat, J. Illiopoulos and Ph. Meyer, 1972），以及格罗斯和贾基夫（1972）提出，如果在模型中包括适量的各种夸克，那么反常将不存在。特别是，他们指出，夸克和轻子之间的反常相互抵消，并且如果模型除了盖尔曼和茨威格[4]原先提出的三种夸克，还包括粲夸克，那么可重正化性就不会被破坏。因此，在建立模型的方法中，由反常论证所得出的关键点在于，仅当夸克以平行于轻子的

方式被包括在模型中时，才有可能建立前后连贯的弱电理论。

弱电理论的特征是存在着有质量中性规范玻色子 Z_μ，它被假定与一个中性的、宇称破坏的弱流相耦合。中性流的预言在 1973 年得到证实。[5]弱电理论的另一个建构，粲夸克，也通过 J/ψ 和 ψ' 粒子的实验发现得到支持[6]，这些粒子等同于粲偶素，同时还得到 D 粒子[7]的实验发现的支持，后者也等同于粲粒子。随着 W 粒子和 Z 粒子在 1983 年的发现[8]，人们认为弱电理论在实验上已得到确认，并且构成了标准模型的核心。

[325] **量子色动力学(QCD)**[9]

在强相互作用的范围内建立模型，是由深度非弹性散射实验中的比约肯标度无关性的经验发现直接确定的，这使夸克模型实体化(并且要求中性胶子作为强子的一部分而存在)，这个发现还要求夸克间的力作为强核力的真正来源，并且强核力仅是夸克间力的残余。在寻找能容纳标度无关性的框架过程中，人们逐渐明白了只有渐近自由理论能解释标度无关性，并且只有非阿贝尔规范理论才可能是渐近自由的(见 10.2 节)。因此物理学家只得在非阿贝尔规范理论的框架内构建强力模型，其中的基本实体不是强子而是夸克和胶子。

他们在建立模型时关键的一步是寻找一个适当的规范群。格罗斯和维尔切克(1973a)在报告他们(重新)发现渐近自由时，提出了一个 SU(3)色规范群。这个主张通过实验数据被有力地提出来了。为了理解这一点，首先要插入介绍一下关于夸克模型中的色量子数。最初，色量子数是作为回应最早的(静态 SU(6)[10])夸克模型所面临的困境而引入的；该模型认为：作为夸克束缚态的最低层的重子属于 SU(6)群的总体上对称的表示，而人们预期夸克出现在 S 态(即在轨道变量上是对称的)。这直接与夸克有 1/2 自旋并因此必须服从费米

统计的假设相矛盾。韩和南部(Moo-Young Han and Nambu, 1965)提出，这个困境可以通过引入一个新附加的三值量子数即色量子数而解决，而重子关于色量子数也将总体上是反对称的。

韩和南部进一步把新对称群和色 SU(3)群或 $SU(3)_C$(除了旧的味 SU(3)群)这种新量子数联系起来，并且为夸克中的"超强相互作用"提出了一个杨—米尔斯类型的理论，这个理论后来被 $SU(3)_C$ 三重态所代替：

> 我们引入八个规范矢量场……其行为……如同 $SU(3)_C$ 中的八重态，而不同于 SU(3)中的单态……形成具有 $SU(3)_C$ 对称性的重子和介子的超强相互作用……最低[重子和介子]质量级别将是 $SU(3)_C$ 单态。
>
> (Han and Nambu, 1965)

虽然这篇论文被广泛阅读和引用，但是只有关于三夸克三重态的思想作为一种建立一致的强子模型的有效方法而得到认真对待，而关于色规范理论($SU(3)_C$ 规范群及其八个色荷规范矢量玻色子、色荷夸克和色中性强子)的"深邃的"思想则被忽视了。在量子色动力学问世之前，这些"深邃的"思想一直未受到赏识，这个事实是完全可以理解的，因为，这些思想在 20 世纪 60 年代中期提供给物理学共同体时，赏识它所需的实验和理论语境，包括标度无关性、重正化群和渐近自由，都还没有出现。 [326]

色量子数的想法在反常论证的语境下，被 20 世纪 70 年代初实验限定极大地强化了。1972 年，有人注意到弱电理论是反常自由和可重正化的，不仅引入包括粲夸克在内的夸克组分是必需的，而且夸克

还必须有三个"色"，以使轻子组分产生的反常可被三色夸克组分所抵消。[11]1973年，巴丁、弗里奇和盖尔曼(Bardeen, Harold Fritzsch, and Gell-Mann, 1973)富有成效地对一种旧的论证作了探索，这种论证表明手征反常的系数能够由低阶微扰理论(三角形曲线图)[12]准确给出，并且在整体手征反常中 $\pi^0 \rightarrow 2\gamma$ 衰变率和强子基本组分的电荷的平方和之间存在一定关系。[12a]他们认为在一定理论语境下，也就是只有当由色量子数提供的 1/3 因子被包括在分数电荷夸克模型中时，这一论证才意味着有可能解释清楚观测到的衰变率。

由于充分意识到关于色量子数的论证越来越多，格罗斯和维尔切克提出核子的中性组分胶子可能负载色量子数[13]并且像色规范玻色子一样起作用：

> 一个特别有吸引力的模型建立在费米子的三个三重态上，它具有作为整体对称性的盖尔曼 SU(3) × SU(3)和提供强相互作用的 SU(3)"色"规范群。也就是说，强相互作用规范群的生成元是和通常的 SU(3) × SU(3)流相对易的，并把具有相同的同位旋和超荷、但有不同"色"的夸克混合起来。在这种模型中，矢量介子是(味)中性的，而电磁流或弱流的算符乘积展开的结构在本质上是自由夸克模型的结构(直到可计算的对数校正)。
>
> (Gross and Wilczek, 1973a)

他们构建强相互作用模型的下一步，是在严格规范对称性和破缺规范对称性之间作出选择。人们可能期望，如同在最初杨—米尔斯理论中的情形，一种未破缺的对称性可能意味着无质量的、强耦合矢量介子的存在。但是，格罗斯和维尔切克指出，在渐近自由理

论中，

> 在(夸克和胶子的)"自由"拉格朗日函数和(强子)态的谱线之间可能存在很小的关联……在此情况下的(夸克和胶子场的)格林函数的红外行为，是由该理论的强耦合极限所决定的。这也许是再好不过了，这种红外行为能抑制除色单态以外的所有态，并且色规范场及夸克都可以在大的欧几里得动量区域中"看到"，但从来也没有作为真正的渐近状态产生过。 [327]
>
> (Gross and Wilczek, 1973b)

在格罗斯和维尔切克的提议中，关键的论证是渐近自由，正是它使其与类似的韩—南部提议不同。通过对照经验约束，渐近自由在短程范围内解释了标度无关性，并提出了一个富有启发性的长程禁闭机制。而且，它也为微扰理论在短程上值得信赖提供了理论基础。因此到1973年末，微扰量子色动力学这种超强相互作用的可计算模型已经初露端倪，虽然它与原先的强核相互作用不是很相关。

注意到下面一点是很有教益的：在1974年7月伦敦高能物理会议上，这样一种协调而似真的理论"多多少少是值得光荣地提及的"(Pais, 1986)。这又是一个认识到科学发现、思想、模型和理论在很大程度上依赖于其语境的例子。[14]

几个月后，随着粲夸克的发现，情况大变。把新发现的J/ψ粒子作为束缚的"粲夸克—反粲夸克"系统或粲偶素，这就意味着粲夸克的质量很大，因此其康普顿波长很小。由于约束夸克的范围比这个长度还小，这就提示了问题只涉及夸克和胶子的短程行为，并可用微扰量子色动力学分析。粲偶素可以被合理地描述为具有禁闭势的

量子色动力学氢原子[15]；这一事实创造了一种新语境，在此语境下，量子色动力学很快为高能物理学共同体所接受。

从1971年特霍夫特的工作使标准模型从地平线上升起，到1975年它在物理学家的基本相互作用的概念中取得了压倒性优势，人们仅花了四年时间就把弱相互作用和电磁作用统一起来，并用可重正化理论加以描述，而强相互作用动力学也从一种混乱状态过渡到可用一种有用的似真模型来加以描述。

11.2 进一步的延伸

虽然标准模型很成功，但是它还包含太多的任意参量，除了弱耦合常数外，还有加进了规范耦合的温伯格角，夸克味的卡比玻角(Cabibbo angles)，以及轻子、夸克、规范玻色子和希格斯粒子的质量参量。因此，从统一方面来说，即从追求规范场纲领的基本动机之一来看，情况还远非令人满意。

[328]

同时，在这些可用的理论工具中，最为重要的是重正化群和对称性破缺，它们已经在推广标准模型和把规范场纲领导向其逻辑结论的过程中，提供了有力的激励作用。尤其是，由乔治、奎因和温伯格(Howard Georgi，Helen Quinn，and Weinberg，1974)用重正化群方程组得出各种能量的强耦合和弱电耦合的计算，表明这两种耦合在相当于大约10^{14} GeV到10^{16} GeV能区附近相等。这种合并，再加上这两个理论有相似的非阿贝尔规范结构的事实，为寻找进一步的统一提供了很强的诱惑，而在大统一理论中，具有不同不变性属性的自然律和对称定律以及反对称物理态，所有这些都源自更高的对称性，而更高的对称性可以用来描述早期宇宙条件下的物理学——当宇宙膨胀、温度降低时，通过一系列相变，而达到用标准模型描述的目前

状态。

在这样一种建立在规范原理基础上的大统一中，物质场(费米子)的内在特性拥有一种动力学表现：规范场和自身耦合，也以唯一和自然的方式跟传递这些特性的物质场的守恒流耦合。因此，量子费米子场形式的物质最终获得其完全成熟意义上的实体地位。不仅物质和力场作为两种实体是统一的，而且物质自身也变得活跃，其活力表现在规范耦合中。

不过，这样一种大统一只有在群 $SU(2)_T \times U(1)_Y$ 和 $SU(3)_C$ 起源于一个更大的单群破缺时才能达到，其中前一种群的规范量子负责相应的弱电力和强力，后一种群的规范量子负责大统一力。除了帕蒂和萨拉姆(Jogesh Pati and Salam, 1973a，b)提出的大统一的最初的 $SU(4) \times SU(4)$ 模型，以及由乔治与格拉肖(1974)提出的 $SU(5)$ 模型之外，还出现过其他几个模型。[16]它们全都建立在同样的基本思想之上，并且有大量的共同特点。

大统一理论(GUTs)的基本思想由泰勒(J. C. Taylor，1976)概括为等级分明的对称破缺一般形式：所有相互作用的最根本的大局域规范对称性，按一系列步骤破缺，给出一个对称性破缺等级体系。从 [329] 一个规范群 G 和希格斯场 ϕ 开始，假定真空期望 $F = F^{(0)} + \varepsilon F^{(1)} + \varepsilon F^{(2)} + \cdots$，其中 ε 是小参量。取 $F^{(s)}$ 的小群为 $G^{(s)}(s = 0, 1, 2, \cdots)$，其中 $G \supset G^{(0)} \supset G^{(1)} \supset G^{(2)}$ 等。那么，G 高强度地破缺到 $G^{(0)}$，在此过程中就产生很重的矢量介子($gF^{(0)}$ 数量级的质量)，$G^{(0)}$ 则强度较低地破缺到具有较轻的矢量介子的 $G^{(1)}$，如此等等。把三种力统一成一种力的优点之一，是任意参量的数目能极大地约化。比如，温伯格角在 $SU(5)$ 模型中能大概确定为 $\sin^2\theta_W = 0.20$。[17] 超重矢量介子的可观测效应在通常的能量下假定很小。

大统一理论的一个共同特点是夸克和轻子合并成相同的多重态。局域规范不变性随后导致具有新性质的新型规范场。比如在最初的SU(5)模型中，有24个统一力场，这些场中有12个量子为已知：光子、两个W玻色子、Z玻色子和8个胶子，剩下的12个是新的，人们给取了一个集体名称X。X玻色子需要维持更大的把夸克和轻子混合起来的规范对称性，因此，它们能把夸克变成轻子或相反。那就是说，X玻色子为重子数不守恒提供了一种机制，其中一个直接推论是质子衰变，这为宇宙学家关于宇宙重子过剩的老问题提供了一个新视角，也即是观测到的物质—反物质非对称性可能有对称的起源。

大统一理论还有许多其他的宇宙学含义，因为这些模型把物理学家的注意力引向早期宇宙，把它作为超高能实验室，其中可创造出质量为$M_x \sim 10^{15}$ GeV的粒子。作为一个理解宇宙结构和宇宙进化的框架，大统一理论虽然在理论上是自洽的，但是却不能产生足够的重子非对称性，以跟观测数据一致，或给出一个有质量中微子，以解决宇宙的“缺失质量”(missing mass)问题。建立在越来越大的统一群上的更加复杂的竞争模型被提出来，提供了新的可能性：新的重夸克和轻子以及大量有质量希格斯玻色子等。由于这些粒子在实验上不可获得，使得不可能在竞争模型之间作出选择。然而，这个事实是否已经暴露了规范场纲领可能超出了其经验基础，还是有争议的。

不管大统一理论是如何成功、如何有广泛性，它们还是没有穷尽规范场纲领的潜力。不久大统一理论就进入了这样一个阶段，在1974年的十一月革命之后不久，规范场纲领成为基础物理学中一个新的正统观念。人们作出更有雄心的努力，以把规范场纲领扩展到

[330] 引力。不过，在讨论这种扩展之前，考察一下物理学家关于规范原理的普适性的态度是适时的。

声称规范原理是普遍可用的说法，在涉及外部对称性时受到一些物理学家的质疑。有两个问题：(i)引力相互作用是由对称性原理即广义协变性原理决定的吗？(ii)我们能否规范一个外部对称性以得到引力量子？

对第一个问题，爱因斯坦最初声称的广义协变性和等效原理一起导致广义相对论，早在1917年就受到克雷奇曼的批评(见4.2节)，后来又受到嘉当、惠勒和弗里德曼等其他人的质疑。[18]这些批评认为广义协变性原理没有物理内容，因为无论什么非协变的理论总能改写成协变的形式。他们声称这种改写唯一要做的只是用协变导数代替通常的导数，增加一个仿射联络 Γ_{ij}^{k} 到理论中，并说存在一个坐标系，其中仿射联络 Γ_{ij}^{k} 的元恰好等于零，而最初的非协变形式系统仍有效。

但有些为爱因斯坦辩护的人却认为，引力相互作用实际上是由局域外部对称性所决定的，因为在非协变理论的重新形式化中，批评者也不得不承认仿射联络 Γ_{ij}^{k} 必定消失。但这就等于是假设时空是平坦的并且外部对称性是整体的。然而，真正的广义协变性是一种局域对称性，这是唯一一种满足引力可纳入其中的弯曲时空的对称类型。因此，这些重新形式化的理论中假定的广义协变性是虚假的。

对第二个问题也并非没有争议。但是规范原理的忠诚追随者作出了许多努力，试过各种不同的对称群，试图找到引力的量子理论。[19]其中，黑尔(F. W. Hehl)及其合作者的工作证明了庞加莱群即洛伦兹变换和平移群可以导出引力的规范理论；若给予积极的解释，这是看来最合情理的引力规范理论。

可惜所有早期试图把引力铸成规范场形式的努力，都未能重正化他们的模型。[20]这种情形让我们想起还未和电磁力统一之前的弱力。

引力和弱力都是不可重正化的。不过在弱力的情况下，一旦找到适当的规范对称性($SU(2)\times U(1)$)，所产生的统一弱电理论的无穷大就会消失。在此教训的指导下，理论物理学家开始寻找比外部对称性存在的群更有效的群，以使引力可重正化。

[331] 把引力和其他力统一起来的困难，部分在于各种理论几乎没有共同之处，甚至是有矛盾的。适合规范引力的对称性，在种类上都不同于规范其他力的对称性。规范引力的对称性都是外部对称性，而所有其他规范对称性都是内部对称性，而且跟时空中任何坐标变换都没有关系。在20世纪60年代中期，有些努力其目标是为了达到外部对称群和内部对称群的统一。[21]不过事实证明，在大量的一般性条件下，庞加莱群(P)和内部对称性群(T)的统一群(G)只能在直积$G = T\otimes P$的平凡意义上得到。人们导出了大量的禁戒定理(no-go theorems)，禁止对称群G以非平凡方式(即不是作为直积)包含T和P，以致对G的这一强制要求有可能允许对具有不同内部对称性的粒子进行统一的描述。[22]

通过引入超对称性，至少可以部分地回避这种禁戒定理。[23]内部对称性只与具有相同自旋的粒子有关，而超对称性通过把非对易数合并为基本元素，把费米子跟玻色子联系起来。其结果是，超对称性的超多重态由一系列不同的内部对称性的多重态组成，尽管超对称性自身正是一种特殊的外部对称性(见下面)，而不是一种外部对称性和内部对称性的真正统一。[24]

为了理解上面这种说法，我们只要回顾一下在量子场论中，玻色子场的维度是1而费米子场维度是3/2，就不难理解联系玻色子和费米子的两个超对称性变换会导致一个统一维度的缺口。[25]和场不一样而又可用来填平缺口的仅有的维度对象是微分。因此在任何整体超

对称性模型中，我们总能在纯粹维度的基础上找到出现在双重变换关系中的微分。因此，在数学上，整体超对称性类似于对变换算符取平方根。那就是为什么它被看作是庞加莱群的扩展(所谓“超庞加莱群”)而不是内部对称性的原因。

在定域理论中，平移算符不同于从点到点的情况。这正是导致一些物理学家认为借此可描述引力的广义坐标变换的概念。的确，在局域超对称不变性的需要和使用“诺特方法”的引导下，有些物理学家得到无质量的、自旋为3/2(引力微子，gravitino)的场规范超对称性和无质量的、自旋为2(引力子，graviton)的场规范时空对称性。因此，超对称性的定域规范理论意味着引力的定域规范理论。这正是为什么定域超对称性理论叫做超引力的原因。

在简单超引力中，超对称性产生子的数目 N 为1。如果 N 大于1，那么理论就叫做扩展超引力。扩展超引力理论的一个特殊性质是，它们除了与引力相关的时空对称性外，还有一个整体 U(N)内部对称性，正是它把相同自旋的所有粒子联系起来，并因此跟超对称性没有什么关系。这证明内部对称性可能是局域的，以致非引力相互作用可以合并进来。在一个 $N = 8$ 的额外局域 SU(8)对称性模型中的发现，极大地激发了某些物理学家；他们希望局域 SU(8)群能产生大统一理论所需要的自旋为1和自旋为1/2的束缚态。[26]不过尽管有这些进展，超对称性和额外内部对称性之间的关系还是不清楚。因此外部对称性和内部对称性的统一仍然没有达到，引力和其他力之间的关系在超引力的语境中仍然不清楚。 [332]

在超对称性中，足够大的对称性能使超引力重正化吗？初步结果是鼓舞人心的。在扩展超引力时表明，由于费米子(引力微子)和玻色子(引力子)之间的对称性，使一次和二次量子修正中的S矩阵抵

消。[27]那就是说，包含引力微子的新图解法消除了由引力子引起的发散。然而，尽管有过早期的乐观主义，物理学家关于超引力的热情很快消退。随着20世纪80年代中期超弦理论(superstring theories)的出现，就很少再有物理学家相信量子引力的问题可在超引力的框架内得到解决。

11.3 规范场纲领：几何纲领与量子场纲领的综合

在第一篇我们考察了几何纲领；按照这种纲领，相互作用可以通过连续经典场实现，而这种场与度规、仿射联络和曲率之类的时空几何结构是密不可分的。这个几何化相互作用的概念导致了对引力的深刻理解，但在物理学的其他地方没有成功。的确，电磁理论和弱相互作用理论以及强核相互作用理论总体上走上了不同方向；在20世纪20年代后期，随着量子电动力学的出现，为基本相互作用开辟了量子场纲领，这一点我们在第二篇已考察过。按照量子场纲领，相互作用通过量子场实现，量子场引起局域耦合和场量子的传播，而跟时空的几何无关。

这两种纲领之间的概念差别是如此之深，以致我们不得不说在这两者之间发生了本体论转变。[333] 这里我们遇上了一种典型的概念革命(conceptual revolution)：两种前后相继的研究纲领有不同的本体论承诺，在此基础上它们具有不同的传递基本相互作用的机制。历史事实虽然如此，但我还是认为仍然能够找到在量子作用和时空的几何结构之间的密切关联，有关证据完全背离了通常对不可通约性(incommensurability)论题的认同。

逻辑上，有三种方法能够建立这种关联，事实上在最近40年中已有物理学家对此作过探讨。首先，人们可以试着在弯曲的时空流

形上、而不是在平坦的闵可夫斯基时空流形上定义量子场。[28]在这种情况下，由于弯曲的时空流形是一个理论在其中得以表述的固定的经典背景结构，因此，新表述不同于传统量子场论的唯一差异，实际上有赖于以不同于传统方式量子化的系统得以建立其上的非闵可夫斯基流形的几何和拓扑结构。然而背景结构的这种固有观点，直接与广义相对论的真正精神相抵触，因为后者的时空结构是自身的动力学实体。不过与我们的讨论更相关的，是在这种表述中的相互作用本质上不是几何的这个事实。

第二，人们可以把时空结构作为量子变量而不是一个不变背景，并使其服从量子力学的动力学定律。在 20 世纪 60 年代，量子化的引力正则方法沿着这条路线导出了著名的惠勒—德威特方程。[29]还有，除了不能成为可重正化表述外，正则量子化还要求把时空分成三个空间维度和一个时间维度，这也与相对论的精神相违背。[30]还有关于等时对易关系含义的解释问题。这些关系只有基于一个固定时空几何上的物质场才能很好定义，但是一旦几何结构被量子化并服从不确定原理，说两点是类空分离的就无意义了。在 20 世纪 70 年代，量子化的路径积分方法变得流行起来，霍金在把路径积分用于四流形 M 上的所有黎曼四度规，并对具有一些所要求边界的每个 M 求和的基础上，形成了他通向量子引力的"欧几里得"方法(1979)，并当流形只有一个简单的相联结的三边界时，得到了一个关于宇宙的"无中生有创世"理论('creation ex nihilo' theory)。[31]但是这个形式化系统也不可重正化。

量子化引力的不断失败使得有些物理学家，比如彭罗斯(1987)，寻求从根本上修正量子理论的结构；其他物理学家，比如超弦理论学 [334] 家[32]，则试图寻求超越量子场论框架的完全不同的物理学理论；更有

像艾沙姆(Chris Isham, 1990)，试图知道通常的量子场论能推广得多远，以及支撑引力纲领的几何和拓扑思想在量子场论中仍能保留的程度，他求助于一种关于时空的几何和拓扑结构的最基本思想，那就是在宇宙中只有有限点和只有有限点集的离散拓扑，它们能解释为时空的结构。虽然超出量子理论、量子场论或时空理论的现存框架也无可非议，但是所有这些回应都还没有产生任何经得住检验的物理理论。由于这些原因，综合几何纲领和量子场纲领的第二条路线没有成功。

第三，这也是这一节的主题，人们总是试图搞清规范理论的几何意义。虽然规范理论的框架只有在引力之外的领域才成功，但是关于它的几何解释还是体现了在本义上的几何纲领和量子场纲领的综合。也正因为如此，量子场纲领本身是规范场纲领的子纲领。

不过为了讨论这个综合问题，我们仅仅需要把注意力集中在使规范场纲领在总体上不同于量子场纲领的基本原理之上，也就是集中在规范原理之上。

规范不变性需要引入规范势来补偿在不同时空点上内部自由度的额外变化。这模仿了广义相对论中把所有可能坐标变换群看作对称群，从而需要引入引力势[33]，用来补偿时空自由度的额外变化。广义相对论中的引力势与特定种类的几何结构关联，这就是四维弯曲时空中的线性联络。规范势也可被解释成与某种类型的结构有关，那就是在纤维丛(fiber bundle)这种时空的一般化结构上的联络。在广义相对论和规范理论之间的这种理论结构上的深刻相似，使得规范理论在本质上是几何的这种可能性得以存在。而且广义相对论本身也可以视为规范理论的特殊情况，这一事实加深了这种印象。因此，杨振宁这位规范场纲领的奠基者之一，已经多次谈到“规范场的几何

本质”，并通过规范场纲领倡导“物理学的几何化”(Yang，1974，1977，1980，1983；Wu and Yang，1975)。他甚至声称规范场正是爱因斯坦为了得到他的统一场论所尽力寻找的几何结构(1980)。

这种说法的出发点当然是引力的几何化，这至少在几何纲领的弱意义上是成立的(见5.2节)。令庞加莱对称性局域化，就等效于广义协变性要求，从而去除了时空的平坦性，并且要求引入非平凡几何结构，比如与引力相关的联络和曲率。 [335]

对其他能被规范势所描述的基本力，人们很容易发现相应理论的理论结构严格地与引力的情况相对应。有一些内部对称性空间：相空间，对电磁学而言就像是环；同位旋，看起来就像三维球的内部；强相互作用的色空间；等等。在每个时空点上定义的内部空间称为纤维(fiber)，而这种内部空间和时空的结合称为纤维丛空间。如果在不同时空点上物理系统的内部空间的方向假设是不同的，那么局域规范对称性就消除了纤维丛空间的“平坦性”，并且要求引入规范势，后者作为规范相互作用的原由，联结了不同时空点上的内在方向。由于规范势在规范理论的纤维丛空间中所起的作用与广义相对论中弯曲时空的仿射联络所起的作用相同，因此似乎证明了杨振宁的几何化论点。

不过，相应的批评会说这不是真正的几何化，因为规范理论中的几何结构仅仅是在纤维丛空间上定义的，而纤维丛空间只是一种数学结构(时空和内部空间的平凡积，两个因子空间彼此没有关系)，跟真正的时空不一样，因此所谓的几何化只能视为无意义的巧辩。

意见上的分歧似乎只是术语学上的，然而进一步的考察揭示了它有更深的含义，并且是理解基础物理学的方法。

为了理解这一点，让我们从一个简单问题开始：几何学是什么？

在狭义相对论之前，唯一的物理几何学是三维空间中的欧几里得几何，其他的几何学虽然在逻辑上是可能的，但是仅仅作为一种精神想象。随狭义相对论而来的是四维时空中的闵可夫斯基几何。今天没有人会否认这是真正的几何。这个信念的重要根据是，定义闵可夫斯基几何的四维时空不仅仅是三维空间和一维时间的平凡积，而且是因为洛伦兹旋转把空间指标和时间指标结合起来了。四维时空中的[336]黎曼几何又不同于闵可夫斯基几何。这也是一种真正的几何，还是仅仅是一种以闵可夫斯基几何为背景描述引力的数学技巧？

广义相对论给我们的一个教训就是引力相互作用必须在时空的几何描述中加以思考。即使我们从闵可夫斯基几何出发，并且在量子场论中重新对广义相对论形式化，其中引力是用无质量自旋为 2 的场描述，最终理论总是涉及弯曲度规，而不是最初假设的平坦度规。[34] 这里平坦度规和弯曲度规之间的关系，与量子场论中裸电荷和着衣电荷之间的关系很类似。因此黎曼几何应当被看作真正的时空几何。或者等效地说，我们应当认识到，真正的时空几何的特征实际上受引力影响。这正是黎曼在其著名的就职演说中所猜测的和爱因斯坦一生所坚持的东西。

那么涉及内部对称性(或等效地说，引力之外的相互作用)时情形又怎么样呢？为了适当评价这种情形，对广义相对论之后的几何观念的进一步演化多讲几句是适当的。

1921 年，爱丁顿引人了一个关于“世界结构的几何，作为空间和时间以及万物的共同基础”的思想。在这一最宽泛的意义上，爱丁顿几何包含了与物质和物质系统的相互作用机制相关的元素。这样一种几何是爱因斯坦和薛定谔在他们的统一场论中所探讨的，也是卡卢察和克莱因在其五维相对论中所探讨的(见 5.3 节)。正如我们将

要看到的，最近二十年来对爱丁顿几何的兴趣已经在规范理论的纤维丛[35]描述中以及在现代卡卢察—克莱因理论中复活了。[36]

在最初的杨—米尔斯理论中，所谓的微分形式体系、内部自由度，虽然是从简单的复数的相推广到了李群的生成元，但是和时空毫不相干。杨振宁和米尔斯仅仅推广了量子电动力学，对其几何意义毫无涉及。但在1967年，内部自由度和外部自由度之间的接触点出现了。吴大峻和杨振宁(1967)提出了同位旋规范场方程的一个解，方程为：$b_{i\alpha} = \sum_{i\alpha\tau} f(r)/r$，其中 $\alpha = 1, 2, 3$，代表同位旋指标，i 和 τ 是时空指标，而 r 是三维矢量(x_1, x_2, x_3)的长度。这个解明确地把同位旋指标和时空指标结合起来。在概念上，这是一个重大的进展。如果闵可夫斯基几何能够因为其时间指标和空间指标与洛伦兹旋转混合起来而被认为是真正的几何，那么吴—杨的解就为新的真正几何提供了一个更宽泛的基础。 [337]

在此方向上的进一步发展是SU(2)规范理论的单极子解，是特霍夫特和波利亚科夫分别在1974年提出的。在这种情况下，单极子解把物理空间中的方向和内部空间中的方向联系了起来。同种类型的联系也出现在规范理论的瞬子解中(参阅10.4节)。

规范理论的几何含义在杨振宁(1974)提出规范场的积分形式体系时已很清楚。积分形式的基本点是引入不可积相因子概念，其定义是 $\phi_A(A + \mathrm{d}x) = I + b\mu_\mu^k(x)X_k\mathrm{d}x^\mu$，其中 $\mu_\mu^k(x)$是规范势的分量，X_k 是规范群的生成元。借助于这个概念，杨振宁发现：(i)列维-齐维塔平移的概念是规范群GL(4)中不可积相因子的特殊情形；(ii)线性联络是规范势的特殊类型；而(iii)黎曼曲率是规范场强的特殊情形。虽然杨振宁具体的GL(4)引力规范理论因为与爱因斯坦的

理论不同，并且可能与观测相矛盾，而受到帕维勒（R. Pavelle，1975）批评，但是杨振宁的一般思想还是得到广泛认可。这个工作令人印象深刻地表明，规范场的概念是深度几何化的。这个印象还因为吴大峻和杨振宁（1975）关于规范场的整体表述工作得以加强。这里规范场的整体几何内涵，包括那些由AB效应和非阿贝尔单极子所蕴涵的东西，都用纤维丛概念得到了探索和表述。[37]

考虑到上述进展在规范理论的纤维丛方案中达到顶点，很明显，与时空自由度相联系的外部几何相平行，存在着与内部自由度相联系的内部几何。这两种类型的几何都能置于同样的数学框架之中。然而，尽管有令人印象深刻的数学相似性和更令人印象深刻的把内部自由度和外部自由度结合在一起的可能性，这两种类型的几何的真正统一（与引力和其他基本相互作用的统一相联系）的绘景，仍然十分模糊。

这个绘景在现代卡卢察—克莱因理论中变得清晰多了。在最初的卡卢察—克莱因理论中，空间的额外维度在低能实验中是紧致化的，是为了与电磁学相一致而被结合到已知的四维时空之中。在复兴了的卡卢察—克莱因理论中，额外空间维数变成7，这是体现在大[338]统一理论中的对称运算数目，并且也考虑了扩展到 $N = 8$ 的超引力。它是假设7个额外空间维度在低能时紧致化为超球。这个七维球包含许多额外的对称性，正是它们力图把力场的基本规范对称性模型化。这意味着内部对称性是与额外紧致空间维度相联系的几何对称性的一种表现，并且所有与内部对称性相联系的几何都是真正的空间几何，也就是与额外空间维度相联系的几何。

然而，在这种情况中，内部几何和外部几何之间的关系是什么样的，这一问题是空洞的。但在深层意义上，这个问题是深刻的。我

们讨论规范场论的几何化时真正的问题是什么呢？ 几何化论题只有在四维时空中的几何结构切实与除了引力或超引力外的规范相互作用相联系时才有意义，或者等效地说，只有把它们跟与额外维度相联系的几何结构混合起来才有意义。 除非达到假设的紧致标度，即普朗克标度 $T = 10^{38}$ GeV 左右，否则我们不能知道情况是否就是如此。因此，直到达到这个标度以前它将始终是个不解之谜。 不过，现代卡卢察—克莱因理论确实通过额外维度的几何结构，打开了建立非引力规范势和四维时空中的几何结构之间的相互关系之门。 然后，在这种理论语境中，几何化论题，也就是与引力和其他规范相互作用的统一相联系而非等同的问题，原则上是可检验的，并且不能被指责为不可证伪的或与基础物理学的未来发展无关。

注意到下面的情况是有趣的：在规范场纲领的几何解释的语境中，区分规范场纲领和量子场纲领的，恰恰是规范场纲领与几何纲领的相似之处。 因此尽管规范场纲领和几何纲领之间的根本区别，正如量子场纲领和几何纲领之间的区别一样基本，然而它们之间深刻的内在关联也很明显。 两者都是场论。 两者都受对称性原理引导，并因此把相互作用与几何结构联系起来。 在从几何纲领到量子场纲领的过渡中，当然，有些基本思想发生了转换：量子场取代了经典场，时空对称性被推广到结合了时空对称性和内部对称性的规范对称性，如此等等。 所有这些转换都是在一个与几何纲领竞争的研究纲领，即量子场纲领中完成的。 因此，如果我们把规范场纲领表述成一种研究纲领，根据这一纲领，相互作用是通过量子化规范场实现的，而量子化规范场是与一些广义几何结构密切相关的，那么认为规范场纲领是几何纲领和量子场纲领的综合，似乎是合理的。

[339] ## 11.4 停滞与新方向：有效场论

在20世纪70年代后期、80年代初期的一小段乐观主义时期之后，正如温伯格指出的，规范场纲领“在对基本粒子的属性的解释或预言上，已经不能进一步取得超出20世纪70年代初期业已很好达到的进展”。在概念上，他认识到，“不必相信量子场论是基础理论，我们也能理解它在TeV左右的低能范围的成功”(1986a, b)。但是，量子场论的困难，甚至在其标准模型的最复杂形式中，从20世纪80年代中期开始就变得越来越明显：量子色动力学对低能π介子—核子相互作用的解释看来几乎难以做到。由于与如下一些问题相关的困难，诸如四维时空中夸克禁闭问题、希格斯粒子的存在问题、代的问题等等，甚至是弱电理论的自洽性似乎也受到怀疑。不仅如此，弱电与强相互作用的统一也遭到攻击，虽然攻击没有成功。更不用说引力的量子化及其与其他相互作用的统一。

同时，大多数像温伯格(1979)那样的目光最锐利的理论学家更清楚地认识到，从20世纪70年代后期开始，在过去的40年中，量子场论的概念基础已经发生了一些根本性的转变，产生了两个结果，一个是试图在理论中解决概念反常，另一个是量子场论和统计物理学之间富有成果的交互作用。对涉及诸如规则化、截止、维度、对称性和可重正化性这些概念的模糊假设得到了澄清，并且对这些概念的最初理解也在转变。诸如“对称性破缺”(自发的或反常的)、“重正化群”、“从低能现象到高能过程中的退耦”、“有意义的不可重正化理论”和“有效场论”这些新概念也在发展，并大量借用在统计物理学中取得的重大进展。这些转变的核心是出现“破缺标度不变性”和与之相关的重正化群方法。

正是温伯格(1978)首次吸纳了这些主要由菲舍尔、卡达诺夫和威

尔逊在批判性的现象学(见8.8节)语境中发展起来的物理学洞见，诸如重正化群方程组不动点解的存在和在耦合常数空间中通过不动点的轨道条件，需要用更基本的所谓“渐近安全”(asymptotic safety)的指导原理解释甚至代替可重正化性原理。不过，这个纲领很快又因 [340] 另一个纲领而黯然失色，即“有效场论”(effective field theory)纲领，这也是由温伯格首创的。一开始，有效场论由于还把可重正化性作为其指导原理，而比渐近安全理论较少雄心。不过，最终它导致了对重正化的新理解，并且对可重正化性的基础性提出巨大挑战，因此澄清了量子场论的理论结构及其本体论基础，并且最重要的是引入了基础物理学前景的根本转变。

根本性转变

截止

正如我们在7.6节看到的，重正化程序由必不可少的两步组成。第一步，针对给定理论如量子电动力学，确定一套算法，把能确定下来的低能过程从未知的高能过程中明确分离出来，后者只有用新理论才能描述。第二步，在理论所描述的低能物理上被忽略的高能过程效应的并入，是通过重新定义理论的有限个参数完成的。重新定义的参数在理论上是不可计算的，但是可由实验确定(Schwinger，1948a，b)。这种并入(incorporation)和重新定义(redefinition)所隐含的要求之所以是可能的，将在下文中考察，这里我们聚焦于分离。

施温格(1948b)和朝永振一郎(1946)直接用接触变换(contact transformations)分离无穷项，未知的贡献简单地由具有适当规范变换性质和洛伦兹变换性质的离散项表示。不过在他们的公式中没有任何有关的线索，能够知道把可知能量与未知能量区域分离开来的边界，它被隐藏在发散积分中的某个地方。因此并入和重新定义只能

被视为发散量必不可少的公式化操作，而在逻辑推理上非常薄弱。[38]

费恩曼、泡利和维拉斯以及大多数其他物理学家，采取了一种不同于施温格和朝永振一郎的方法。他们借助于规则化程序暂时修正了理论，以便使积分有限。在费恩曼(1948c)以及泡利和维拉斯(1949)引入的动量截止规则化方案中，从未知区域中分离出可知区域的分界线，是通过引入的动量截止清楚地指示出来的。[39] 低于截止，[341] 假定该理论是值得信赖的，并且高阶积分的修正可以正常进行和计算。产生在高于截止区域的未知高能过程，不得不从考虑中排除出去。从这一点来看，费恩曼的方案在弥补重正化的基本思想方面，似乎优于施温格首次清楚阐明的施温格方案，这似乎也是逻辑上和数学上更可取的。

然而，各种规则化方案必须回答如下难题：如何把从低能现象中排除出去的高能效应考虑在内？这个问题在定域场论中是很具体的，并且在那个框架内是不可避免的。后来获得正统地位的费恩曼解，是要在计算结束时把截止取为无穷大。按照这种方法，所有高能过程都得到了考虑，并且它们对低能现象的效应也可通过重新定义参量合并进来；在施温格的方案中，重新定义参量出现在所指定的理论的拉格朗日函数中。为此付出的代价是截止不再被当作阈能；在阈能处，理论不再成立，并且为了正确地进行物理描述，要求有新的理论。否则，一系列概念反常就会出现：把截止取作无穷大意味着，理论在每处都值得信赖，而高能过程不是不可知的。这与重正化的基本思想直接相抵触，并且把截止取作无穷大时导致的发散积分清楚地表明情况并非如此。[40]

把截止取作无穷大有重大意义。首先，把可知区域和不可知区域分开的边界线变得深藏不露了。也改变了截止的地位，即把截止

从一种暂时性的阈能变成一个纯粹形式化的措施，并因此使费恩曼—泡利—维拉斯方案还原为施温格的纯形式化方案。在此情况下，人们能够把取代施温格正则变换的费恩曼动量截止规则化，变成一个控制发散量的更有效的形式化算法。或者等价地说，人们能够把施温格对发散积分的直接处理，看作是费恩曼先引入有限截止，然后再把截止取为无穷大这样两步的联合。更重要的是，把截止取为无穷大这一步也加强了通常的形式主义说法，即认为物理学应当独立于截止，从而所有明确指示的截止都应该转移到重新定义参数上来。这种说法似乎是迫于形势，因为这一步确实剥夺了依赖于截止的量的任何物理意义。相反，这种说法反过来允许人们对截止采取一种形式主义解释，并迫使截止从物理学中去除。

如果采取严格的截止，并实在地把它解释成新物理学的阈能，情况会怎样呢？那么，正统的形式主义方案就会失败，而且整个情况会改变：不可能将截止取到无穷大，并且与此相反，物理学也不能声称自己不依赖于截止。事实上，自 20 世纪 70 年代中期以来，对重正化的物理学上和哲学上的理解的重大进展，正是来自于这样一种实在论解释。[41]导致这种立场的物理学推理的线索，是相互缠绕在一起的。因此，为了澄清围绕这个问题的概念情况，人们不得不对它们进行梳理。 [342]

作为开始，让我们考察一下可以对截止采取实在论立场的原因。正如我们前面所注意到的，把截止取作无穷大的动机是为了在低能现象上把高能效应考虑进去，高能效应是通过引入有限截止而被排除出去的。如果我们能找到其他保留这些效应同时又使截止有限的方法，那么让截止无穷大就毫无理由了。事实上，实在论立场从 20 世纪 70 年代末以来变得有吸引力，正是因为理论学家逐渐认识到，不

把截止取作无穷大，高能效应也能保留。通过增加有限数量的新的不可重正化相互作用(这些相互作用和原初的拉格朗日函数具有相同的对称性)，并结合对理论参量的重新定义，也能最终达到相同的目标。[42]应该注意到，不可重正化相互作用的引入不会引起任何困难，因为理论有一个有限截止。

采用实在论立场是要付出代价的。首先，形式体系在增加新的补偿性相互作用后变得更加复杂。然而，只存在着有限数量的需要增加的新相互作用，因为这些相互作用都服从各种限制。而且，实在论立场在概念上比形式主义立场更简单，因此代价不是很高。其次，实在论的形式体系假定只对截止能量有效。由于任何实验都只能探测有限范围的能量，因此，实在论形式体系的这个极限实际上不会在精确度上引起任何真正的损失。因此，表面上的代价只是一种错觉。

下一个问题是如何清楚地表述截止的物理实现，以便找到确定其能量标度的方法。实在论理论中的截止不再是一个形式主义方法或任意参量，而是作为量子场论等级结构的体现而获得其物理意义，并且作为一个分离能量区域的界线，能通过不同参量集和有不同对称性的不同物理定律(相互作用)单独描述。[343] 自发对称性破缺的发现和退耦定理(见下文)提醒人们，截止的值是把与自发对称性破缺有关的重玻色子质量联结起来。既然对称性破缺可以使一些微不足道的不可重正化相互作用成为可探测的[43](原因是其他所有相互作用由于对称性的禁戒而缺位)，因此截止的能量标度能够通过测量理论中不可重正化相互作用的强度而建立起来。

上述初步论证已经表明，截止的实在论概念并非站不住脚。然而，对它的可行性的可信证明，只有把它结合进新的概念网络时才是

可能的，这种新网络为理解可重正化性、不可重正化相互作用和一般量子场论提供了新基础。让我们转向这个网络中的其他线索。

对称性和对称性破缺

在传统程序中，把解的无效(发散)部分从有效(有限)部分分离出来之后，高能效应就被吸收到修正后的理论参数中。然而，为了使这种结合成为可能，模拟不可知高能动力学的振幅结构必须和低能过程中的一样，否则多重重正化就会成为不可能。为了确保所要求的结构相似性，必须作出一个重要假定，事实上该假定是隐含地融入多重重正化的方案本身中的。这就是假定高能动力学被那些限制低能动力学的同样的对称性所限制。由于理论的解构成了理论的对称群的表示，因此如果不同的对称性由不同能区的动力学显示，那么这将隐含着不同的群论限制，进而也隐含着，在不同的动力学部分中的解有不同的结构。如果情况果真如此，那么理论的可重正化性就会遭破坏。

在量子电动力学的情形中，可重正化性是由有点神秘的U(1)规范对称性的普适性所确保的。不过，随着对称性破缺[20世纪60年代初的自发对称性破缺和60年代末反常对称性破缺(ASB)]的发现，情况变得越来越复杂[44]，从而要求对上述关于对称性和可重正化性之间的关系的一般考虑必须予以精炼。

首先来看自发对称性破缺。在凝聚态和统计物理学中，自发对称性破缺是一个涉及动力学系统解的属性的说法，即在能量上有些反对称组态比对称组态更为稳定。本质上，自发对称性破缺涉及解的 [344] 低能行为，并且断言有些低能解表现出的对称性，比由系统的拉格朗日函数表现出的对称性更低，同时其他解则拥有系统的全部对称性。追根到底，自发对称性破缺是系统的内在属性，因为反对称解的存在

和确定是由系统的动力学和参数完全决定的。它们跟解的等级结构密切相关，这在统计物理学中表现在连续(二级)相变现象之中。

在量子场论中，自发对称性破缺只在涉及连续对称性的规范理论中才有物理意义。否则，其数学预言之一(无质量戈德斯通玻色子的存在)就会和物理观测相矛盾。在规范理论的框架之中，上面所提到的涉及自发对称性破缺的所有陈述都是有效的。另外，还存在另一个跟我们的讨论相关的重要论断。在规范理论中，与明确的对称性破缺的情况形成对比的是，各种低能现象借助于自发对称性破缺都能纳入一个等级体系，而不会破坏理论的可重正化性(见 10.3 节)。其中的原因是，只有在能量低于对称性破缺的标度时，自发对称性破缺才会影响物理学结构，因此对于理论的高能行为，从本质上说，不会影响理论的可重正化性。

一般来讲，反常对称性破缺是由量子力学效应引起的经典对称性的破缺(见 8.7 节)。系统在其经典表述中所拥有的某些对称性，有可能在其量子化方案中消失，因为后者可能引入一些违反对称性的过程。在量子场论中，这些情况会因圈图修正* 而出现，同时也跟重正化程序以及不变调节子的缺位有关。

反常对称性破缺在量子场论中起着重要作用。特别是，对保证对称性不反常破缺的渴望，会给模型的建构施加一个很强的限制(见 11.1 节)。如果涉及的对称性是局域的，比如规范对称性和广义协变性，那么注定会出现反常对称性破缺，因为理论的可重正化性已被破坏。如果所涉及的对称性是整体的，那么反常对称性破缺的出现

---

* 在量子场论中，圈图的贡献是无穷大，为了避免发散困难，唯象计算中将它忽略掉。——译者

就是无害的，或者甚至是人们所希望的，例如在解释 $\pi^0 \to \gamma\gamma$ 衰变的整体 γ_5 不变性的情形中，或者为了获得作为束缚态的有质量强子，在具有无质量夸克的量子色动力学中标度不变性的情形中。 标度反常的最深刻含义无疑是重正化群的概念，不过这一点我们已在 8.8 节评论过。

退耦定理和有效场论(EFT) [345]

按照重正化群方法，不同的重正化规定只会导致理论的不同参数化。 退耦定理体现了在选择恰当规定中这个自由度的重要应用；这个定理首先是由西曼齐克(1973)、然后是阿佩尔奎斯特和卡拉宗(T. Appelquist and J. Carazzone， 1975)加以表述的。 这个定理建立在方次计数论证基础上，并且在涉及其中有些场有比其他场更大质量的可重正化理论时，该定理断言，在这些理论中总能找到重正化规定，使得重粒子能被证明可从低能物理状态退耦，但会产生被重质量所划分的实验动量的幂所抑制的重正化效应和修正。

这个定理的一个重要推论是低能过程可由有效场论描述，而有效场论只吸纳那些在所研究的能量方面有实质重要性的粒子。 也就是说，没有必要去解那个描述所有轻重粒子的完备理论(Weinberg, 1980b)。

借助于重正化群方程组、耦合常数、质量、格林函数的阶，通过从完备的可重正化理论删去所有重场以及适当地重新定义，可得到一个有效场论。 显然，有效场论对物理过程的描述依赖于语境。 这是由可达到的实验能量限定了的，并因此能够紧密追踪实验的状况。有效场论的语境依赖性，也体现在由与自发对称性破缺相联系的重质量所表示的有效截止之中。 因此，随同退耦定理和有效场论的概念，出现了由量子场论提供的自然等级概念，这解释了为什么在任何

一个层次上的物理学描述都是稳定的，并且不受在高能区所发生的任何事情的干扰，从而为这种描述的应用作了辩护。

在重正化群概念背后的思想和在有效场论概念背后的思想之间，似乎有明显的矛盾。前者是以所考虑的系统中特征标度的缺乏为基础，而后者严格地将重粒子的质量标度作为物理截止或区分有效场论的不同有效区域的特征标度。不过，如果我们记得重粒子仍然对有效场论中的重正化效应作出贡献，那么这种矛盾马上就消失了。因此在有效场论中，重粒子的质量标度只能看作赝特征标度，而不是真正的特征标度。虽然这种赝特征标度的存在反映了在不同能量标度[346]上耦合的等级顺序，但是它也未改变量子场论所描述的系统的基本特征，即特征标度的缺乏和在不同能量标度上涨落的耦合(见8.8节)。高能标度和低能标度上的涨落之间的某些耦合普遍存在，并且在低能物理学的重正化效应中显现出来，而另外一些却被抑制住并且在低能物理学中没有可观测的线索。

“退耦不是绝对的”上述论断得到一个重要观测的强化，这就是重粒子对于低能物理学的影响在有些情况下是可直接探测的。如果在涉及导致这些过程对称性破缺的相互作用的重粒子(比如在弱相互作用中的W和Z玻色子)缺乏的情况下，存在一些恰恰为对称性(比如宇称、奇异数守恒等)所禁戒的过程(比如弱相互作用过程)，那么重粒子对低能现象的影响就是可观测的，即使由于退耦定理，这种影响为被大质量分开的能量方次给抑制住了，情况也是这样。这些效应典型地可用有效的不可重正化理论(比如关于弱相互作用的费米理论)描述，它们是作为可重正化理论(如电弱理论)的低能近似，拥有一个物理截止或由重粒子设置的特征能量标度(比如费米理论的300 GeV)。一旦实验能量达到截止能量，不可重正化理论就变得不能

用，而新物理学看起来要求其描述要么是可重正化理论，要么是具有更高截止能量的新有效理论。第一种选择代表了传统。第二种选择代表对可重正化性原理的基础地位的严峻挑战，并在目前得到发展契机和普及。

对可重正化性的挑战

有效场论的概念阐明了量子场论在不同标度上如何采取不同的形式，并允许有两种不同的看待情况的方法。第一种，如果在高能时可重正化理论是可得到的，那么任何低能的有效理论也能通过对理论的重场求积分的系统化方式获得。因此，可重正化弱电理论和量子色动力学，被理解为某种大统一理论的低能有效理论，已经丧失了假定其为基础理论的地位。另一种可能性也与看待这种情况的方式相一致，即假设存在一个包含不可重正化相互作用的有效理论的塔式结构，每一项比后面的项都有较少的粒子，并且具有较小的不可重正化相互作用。当物理截止(重粒子质量 M)比实验能量 E 大得多时，有效理论是近似可重正化的，因为不可重正化项被 E/M 的幂所抑制。 [347]

第二种方法更接近于高能理论学家在其研究中的实际情况。由于没有人知道可重正化理论在不可达到的高能会是什么样子，甚或它是否存在都未可知，因此他们不得不首先探测可得到的低能区，然后设计符合这个能量范围的表示方法。只有当高能与物理学的理解相关时，他们才把理论扩展到高能区。这个实践过程体现在理论的无尽的塔式概念之中[45]，其中每个理论对特定实验情况都有特定的作用，而且没有一个理论能完全作为基本的理论。按照这种观点，可重正化性的要求被有效场论中不可重正化相互作用的条件所代替：在一个标度 m 上描述物理学的有效理论中的所有不可重正化相互作用，必定由质量标度为 $M(\gg m)$ 的重粒子产生，并因此被 m/M 的幂

所抑制。而且，在包括质量为 M 的重粒子的可重正化有效理论中，这些不可重正化相互作用必定消失。

这些澄清与重正化群方程一起，已经帮助物理学家产生一种对重正化的新理解。正如格罗斯(1985)所说："重正化是一种关于物理相互作用的结构随着被探测现象标度的改变而变化的表述。"注意，这个新理解和旧理解很不同，它完全聚焦于高能行为和防止发散的办法。它表明一种涉及各种物理相互作用随能量标度的有限改变而变化的更一般情况，并因此为考虑不可重正化相互作用提供了足够的余地。

在关于理论建构应当采取哪些东西作为指导原则的问题上，物理学家的态度近年来在有效场论语境中发生了重大改变。多年来，可重正化性已经作为接受一个理论的必要条件。随着人们意识到实验只能在有限能量范围内进行探测，有效场论于是成为许多物理学家分析实验结果的自然框架。由于不可重正化相互作用很自然地发生在这个框架之内，在建构理论模型来描述目前可达到的物理学时，就没有先验的理由把它们排除出去。

除了自身协调一致以及与对重正化的新理解相容之外，其他一些[348] 论证也支持严肃对待不可重正化相互作用。首先，不可重正化理论完全能够适应实验和观测，尤其是在引力领域。第二，它们拥有预言能力，并且还能够通过取越来越高的截止来改进这种能力。第三，因为它们的现象学本性，它们在概念上比可重正化理论简单，正如施温格强调的，这在物理学上引发了对物理粒子的动力学结构的大量质疑。第四个支持论证源自构造论理论学家，自从 20 世纪 70 年代中期以来，他们帮助理解不可重正化理论的结构，并且帮助发现了使不可重正化理论有意义的条件(见 8.4 节)。

反对不可重正化理论的传统论据是，它们在其能量高于物理截止时是不可定义的。 因此让我们来看看不可重正化相互作用的高能行为，许多物理学家已经把它作为量子场论中最基本的问题之一。

在有效场论的最初框架中，不可重正化理论只是作为辅助手段。当实验上可达到的能量接近截止，并且新物理学开始出现时，它们就变得不正确了且必须被可重正化理论所替代。 在温伯格渐近安全理论的框架中，不可重正化理论获得了更为基础的地位。 不过，它们和有效场论仍有共同特点，也就是说，所有关于它们的讨论都是建立在让截止达到无穷大的基础之上，并因此属于截止的形式主义解释的范畴。

然而，如果我们从基本的有效场论思想引出其逻辑结论，那么有效场论的前景就会发生根本变化，就会出现新视角，能发展出一个新的量子场论解释，并且出现了有待于探索的量子场论的新理论结构。有效场论的彻底贯彻，就像乔治(1989b)和勒帕热(Lepage, 1989)证明了的，当实验上可获得的能量接近不可重正化有效理论的截止时，它总能用另一个具有更高截止的不可重正化有效理论来替代。 这样，在截止之上的不可重正化相互作用的高能行为，就能适当地用如下方法处理：(i)由截止的改变和重正化群方程的可计算性引起的重正化效应的变化，以及(ii)附加的不可重正化抵消项。[46]

因此，在发展的任何阶段，截止总是有限的，并能得出一个实在论解释。 除了有限的截止之外，还有两个新的为量子场论的传统结构所缺乏或禁止的、而在有效场论的理论结构中变得合法和不可或缺的部分。 这两部分就是：(i)随截止的具体改变而产生的重正化效应的变化，以及(ii)由有限截止的引入而合法化的不可重正化抵消项。 [349]

应当提及伴随新的重正化概念出现的一些困难。 第一，其初始

假定认为重正化群方程有不动点解，但是不能保证这些不动点总是存在。因此整座大厦的基础并不牢固。第二，有效场论是由退耦定理辩护的。然而，这个定理在对称性破缺时遇到大量的复杂困难，不管涉及的是自发破缺还是反常破缺，这两者都是现代场论模型必然要涉及的。[47]第三，退耦论证没有处理小量假设，即结合在重正化效应中的发散(这也存在于退耦的情况中)实际上是小效应。最后，由朗道首先提出的定域场论的长期存在的困难，即零荷论证(zero-charge argument)，也是至今尚未解决的问题。

原子论和多元主义

在量子场论框架内发展起来的用于描述亚原子世界的各种模型，在本质上仍然是原子论的。出现在拉格朗日函数中用场描述的粒子，终将会被认为是世界的基本组分。在某种意义上，有效场论方法进一步扩展了原子论范式，因为在此框架下所研究的领域更好分辨，并且更明确地定义了等级结构。等级由与自发对称性破缺链有联系的质量标度所确定，并且由退耦定理所辩护。

退耦定理并不反对不同等级层次之间因果联结的普遍思想。事实上，人们总是假定存在着这样的因果联结——最值得注意的是，它们通过在低能现象基础上的高能过程的重正化效应把自身显示出来——并由重正化群方程描述；它们因此构成了这一定理的概念基础。这是试图给出因果联结的普遍意义，并且给出与被否定的科学探究直接相关的规定。更准确地讲，所否定的只是如下假定，即认为有可能仅仅通过这类因果联结，从没有任何经验资料输入的高能标度上的简单性，推断出在低能标度上所出现的复杂性和新颖性。正如退耦定理和有效场论所要求的，有必要把经验资料输入到在低能标度上可应用的理论本体，这种本体在高能标度上跟科学研究没有直接

相关性。这正孕育着一种对物理世界的特殊表示。在此图景中，高 [350]
能区可以分层地看作准自律区域，每一层都有其自己的本体和相关的"基本"定律。每一层的本体和动力学都是准稳定的，几乎不受其他层的任何影响，因此叫做准自律区域。

对建立在有效场论基础上的等级结构的检验，从形而上学的角度产生了两个似乎矛盾的含义。一方面，这种结构似乎倾向于支持用一种还原论甚或重构主义的方式解释物理现象的可能性，至少在自发对称性破缺发生作用的限度内是这样。在过去 20 年中，高能物理学界主流的大多数努力都投入在从标准模型到超弦理论的处理之中，可以看作对这种潜力的探索。因此，这样一种等级结构在弱意义上还是落入原子论的范畴。

另一方面，认真看待退耦定理和有效场论就会认同客观的涌现属性的存在，这意味着可能的理论本体的多元主义观点。[48]这反过来对还原主义方法论设置了一个内在限制。因此，作为原子主义的还原论追求的量子场论的进展，受到其内在逻辑的支配，已经达到一个关键时刻；具讽刺意味的是，在这关键时刻其自身的还原论基础已经遭到某种程度的损坏。

涉及不同层次之间关系时，正是强烈的反还原论信仰，使有效场论培育起来的原子论的多元主义版本不同于传统量子场论采取的粗糙原子论版本，而量子场论的构成部分是还原论和重构主义。另外，强调经验信息这种历史上的偶然成分，也明显把原子论的等级多元主义版本，和新柏拉图主义的数学原子论相对照。在量子场论学家的传统追求中，新柏拉图主义总是被暗中假设为得到正统信仰的验证，这种信仰认为应该将没有历史背景的数学实体看作他们研究的本体论基础，从这种本体论基础中可推导出经验现象。

通向量子场论基础的三条进路

在这一节中所评论的基础的变换以及相关概念的发展，为施温格的远见卓识被接受和得到进一步发展提供了丰厚的土壤。施温格的远见卓识是在其对量子场论的算符表述和重正化的批评中得到发展的，并且在其源论的表述中作出详细叙述。施温格的观点强烈地影[351]响了温伯格关于通向手征动力学的现象学拉格朗日函数方法和有效场论方面的工作。我们很容易找到施温格的源论和有效场论的三个共同特点：第一，它们对基础理论的反叛；第二，它们能够把新粒子和新相互作用结合到现存方案中的灵活性；第三，它们中的每一个都有考虑不可重正化相互作用的能力。

不过，这两个方案存在一个根本区别。有效场论是定域算符场论，并且不包括任何特征标度，因此不得不处理来自任意高能量上涨落的影响。在各种动量标度上的定域场中的局域耦合行为能用重正化群方程追踪，虽然这种方程不是一直有但是常常有一个不动点。如果情况果真如此，那么定域场论就是可计算的，并且能作出有效预言。因此对有效场论来说，有效定域算符场的概念是可接受的。相反，这个概念在施温格理论中总体上是遭到反对的。施温格理论是一个彻底的现象学理论，其中数字场和算符场不一样，只在低能时对单粒子激发起作用。因此，在施温格理论中没有重正化问题。相反，正如我们上面注意到的，在有效场论的表述中，重正化已经采取了越来越复杂的形式，并且已经变成一个更加有效的计算工具。

如果我们严肃地把有效场论看作提出了一幅新的世界图景、一个关于量子场论基础的新概念，那么上面陈述所提到的一些概念困难，就不可能是能用已经建立起来的方法论来解决的常规困难。在处理这些概念问题时所要求的，似乎是我们关于基础物理学自身概念的巨

大改变，即从关注基础理论(作为物理学的基础)，变为关注拥有在各种能量标度上都能发挥作用的有效理论。

许多理论学家反对对有效场论的这种解释。对格罗斯(1985)和温伯格(1995)而言，有效场论只是更深层理论的低能近似，并能用一种系统方法从中得到。然而，值得注意的很有趣的一点是，虽然他们相信在还原主义方法论之内，借助于更复杂的数学或新的物理思想，或者直接通过挖掘亚夸克物理学中越来越深的层次，摆脱高能物理学面临的概念困难的方法迟早会被找到，但是他们两人对作为物理学基础的量子场论都失去了信心，并认为更深层理论或终极理论并不是场论而是弦论，虽然后者在现阶段还不能视为正当的物理学理论。

因此出现了一个具有深远意义和探索价值的问题，这就是：从弦论学家的观点来看，在量子场论的基础中，是哪种缺陷剥夺了量子场论作为物理学基础的地位？对于有些更“保守”的数理学家，比如怀特曼(1992)和贾菲(1995)等等许多物理学家而言，这个问题不存在，因为他们相信，借助越来越多的数学技巧，量子场论的连贯表述(最可能取规范理论这种形式)，能被确认下来并继续作为物理学的基础。 [352]

因此，关于量子场论的基础问题，关于量子场论作为物理学基础的问题，以及为什么不能再把量子场论看作物理学基础的问题，目前实质上有三条进路。除了物理学研究之外，对这三条进路的评价也要求哲学上对还原论和涌现论进行澄清，不过这已经超出本书范围。[49]

注释

1. 关于强子的连贯图景，也需要一些中性成分(胶子，后来被证实就是夸克内力的规范玻色子)。关于理论分析，见 Callan and Gross(1968)，Gross and Llewelyn-Smith(1969)以及 Bjorken and Paschos(1969，1970)。关于实验报告，见 Bloom *et al*.(1969)和 Breidenbach *et al*.(1969)。

2. 派斯提到的大会报告起草人的谈话，能在 *Proceedings of 1968 International Conference on High Energy Physics*（维也纳）以及 *Proceedings of 1970 International Conference on High Energy Physics*（基辅）中找到。

3. 见 Veltman(1992)。

4. 第四种夸克由哈拉(Y. Hara, 1964)首次引入，用以实现弱相互作用和电磁相互作用方面轻子—重子的基本对称性，并且是为了避免所不希望的中性奇异变化流。它成为熟知的粲夸克要归功于比约肯和格拉肖的论文(1964)，并且因为格拉肖、伊奥普洛和马亚尼的论文(Glashow, Illiopolos and Maiani, 1970)而闻名。

5. Hasert *et al*.(1973a, b).

6. Aubert *et al*.(1974), Augustin *et al*.(1974)以及 Abrams(1974)。

7. Goldhaber *et al*.(1976)以及 Peruzzi *et al*.(1976)。

8. Arnison *et al*.(1983a, b)以及 Bagnaia *et al*.(1983)。

9. 量子色动力学这个名称首次出现在 Marciano and Pagels(1978)，并且要归功于盖尔曼。

10. SU(6)群包含一个变换固定自旋的夸克类(或味)的 SU(3)子群，和一个变换固定类(或味)的夸克自旋的 SU(2)子群，因为人们假定夸克形成了味 SU(3)三重态和自旋双重态。

11. Bouchiat, Illiopoulos, and Meyer(1972)；以及 Gross and Jackiw(1972)。

12. Adler and Bardeen(1969).

12a. 见 Adler(1969)。

13. 他们注意到，由于深度非弹性实验表明核子的带荷部分是夸克，中性的胶子就不得不是色中性的。这并不是一个关于夸克是非色中性的论证，而是说，不然的话，胶子在动力学总体上是虚设的。

14. 量子色动力学领域中的另一个例子是由韩—南部的提议和特霍夫特对非阿贝尔规[353]范理论的渐近自由的发现提供的；在弱电理论的领域中，这样的例子由格拉肖—温伯格—萨拉姆模型提供。

15. Appelquist and Politzer(1975), Harrington, Park, and Yildiz(1975), Eichten, Gottfried, Kinoshita, Koght, Lane, and Yan(1975).

16. 例如，见 Fritzsch and Minkowski(1975)；Cürsey and Sikivie(1976)；Gell-Mann, Ramond and Slansky(1978)。

17. Weinberg(1980a).

18. 见 Misner, Thorne, and Wheeler(1973)以及 Friedman(1983)。

19. 例如，见 Utiyama(1956)和 Yang(1974)。

20. 参阅第 10 章的注释 49。

21. 通过与维格纳(1937)提出的核物理学中的 SU(4)对称性(它是自旋旋转和同位旋旋转的联合)进行类比，崎田文二(B. Sakita, 1964)以及居尔塞伊和拉迪卡蒂(F. Gürsey and L. A. Radicati, 1964)提出了把内部对称群 SU(3)和自旋群 SU(2)作为子群包含在内的夸克模型的非相对论性静态 SU(6)对称性。由于自旋是庞加莱群的两个守恒量之一(另一个是质量)，这就被看作外部对称性和内部对称性的一种统一。

22. 例如，见 McGlinn(1964), Coleman(1965)以及 O'Raifeartaigh(1965)。有关评论，见 Pais(1966)。

23. 作为整体对称性的超对称性思想首次出现在雷蒙德(Ramond, 1971a, b)以及内沃和施瓦茨(Neveu and Schwarz, 1971a, b)关于 S 矩阵理论传统中的双重模型的工作中。这个思想随后被韦斯和祖米诺扩展到量子场论(1974)。其最惊人的性质之一是两个超对称性变换导致时空转变。

24. 参阅 van Nieuwenhuizen(1981)。

25. 这个约定源自作用量是没有量纲(在 $h = c = 1$ 的单位制中)的这一要求，还有另一个约定是玻色子场在作用量中有两个导数，而费米子场只有一个导数。

26. 比如，霍金在其就职演讲(1980)中取 $N = 8$ 超引力作为规范场纲领的制高点，或者甚至是理论物理学本身的制高点，因为它原则上能解释物理世界中所有力和所有粒子的

任何事情。

27. 无穷大只在扩展了的超引力理论中抵消，因为只有在那个理论中，所有粒子才都能转化成引力子，并且所有图形才都能约化到只有引力子的情形，从而能证明引力子有一个有限总量。

28. 例如，见 Friedlander(1976)。

29. Wheeler(1964b)，DeWitt(1967c).

30. 如果有人反对外部时间的思想，并试图把时间等同于度规变量的某个函数，然后就这个"内部时间"重新解释动力学，那么最后他会发现实践中不可能做到这种精确的等同。参阅 Isham(1990)。

31. Hartle and Hawking(1983).

32. Green，Schwarz and Witten(1987).

33. 为了比较起见，我们称第二类 $\Gamma^{\alpha}_{\mu\nu}$ 引力势为克里斯托费尔符号。$\Gamma^{\alpha}_{\mu\nu}$ 是用 $g_{\mu\nu}$ 及其一阶导数表达的，$g_{\mu\nu}$ 常被泛泛地叫做引力场。

34. 这被德塞尔(Stanley Deser，1970)证明过。

35. Yang(1974)，Wu and Yang(1975)，以及 Daniel and Viallet(1980)。

36. 关于现代卡卢察—克莱因理论，见 Hosotani(1983)。超弦理论学家也对爱丁顿几何感兴趣，但是超弦理论超出本书范围。

37. R^4 上的纤维丛是时空和内部空间的积空间(product space)的推广，这允许在丛空间(bundle space)中的可能结合，并因此产生时空和内部空间的非平凡合并。数学上，外部指标和内部指标能够通过特定的传递函数混合起来，这些函数在规范理论中是通过广义相位变换起作用的。关于方法，见 Daniel and Viallet(1980)。 [354]

38. 关于狄拉克的批评，见他的论文(1969a，b)。

39. 至于其他规则化方案，同样的说法也适用于格点截止方案，这在本质上是等效的，但不适合维数正规化，它更加形式化并且与这里的讨论无关。

40. 勒帕热(1989)没有进一步解释就断言："这最后一步[取截止至无穷大]在包括量子电动力学在内的许多理论的非微扰分析中看起来好像也是错误的一步。"

41. Polchinski(1984)和 Lepage(1989)。

42. 见 Wilson(1983)，Symanzik(1983)，Polchinski(1984)，尤其是 Lepage(1989)。

43. 这种不可重正化相互作用看上去像是不可达到的高能动力学的低能证据，并因此受由实验能量除以重玻色子质量所体现的能力的抑制。

44. 明确的对称性破缺，比如，增加非规范不变质量项到纯杨—米尔斯理论中，与我们这里的讨论无关。

45. 有效场论的层次可能是无穷的，这一点通过量子场论的局域算符表述得到说明。见 8.1 节。

46. 见 Lepage(1989)。

47. 关于这个论题有范围广泛的讨论。例如，见 Veltman(1977)和 Collins，Wilczek，and Zee(1978)。其主要论证是这样的：如果存在一个在重粒子缺乏时被禁戒的过程，那么源于重粒子的不可重正化效应将是可探测的。

48. 波普尔(1970)令人信服地论证了对理论本体论中的多元主义而言涌现观点的含义。

49. Cao and Schweber(1993).

第 12 章
本体论综合与科学实在论

[355] 前面几章对于20世纪场论的历史研究，为科学如何发展的模型提供了一个充分的检验基础。在此基础上，我在这一章中将论证完成概念革命的一种可能道路，我称之为“本体论综合”（ontological synthesis），从而为科学实在论和科学增长的合理性提出一个论证。

12.1 科学如何发展的两种观点

在现代科学哲学中，关于科学如何发展的问题有许多种观点，我将详细考虑其中两种。按照第一种观点，科学的演化是通过把过去的成果逐渐结合到目前的理论中而取得的，或简言之，科学是持续进步的。经验主义哲学家内格尔(Ernest Nagel)就持这样一种“通过吸收实现增长”的观点。内格尔理所当然地认为知识总是在增加，并声称“相对自律的理论逐渐被某个理论包含或还原的现象是不可否认的，并且是现代科学史中反复出现的特点”(1961)。因此他说科学增长中有稳定内容和连续性，并且把这个稳定内容看作比较科学理论

的共同尺度。可通约性(commensurability)的思想被看作合理地比较科学理论的基础。

"通过吸收实现增长"观的一种更复杂的版本是由塞拉斯(Wilfrid Sellars，1965)和波斯特(1971)提出的。尤其是波斯特诉诸所谓的"广义对应原理"；根据这个原理，"任何可接受的新理论L，在那些先前理论S已被很好地检验的条件下，能通过'退化'成先前理论(S)来解释其先前理论(S)的成功"(同上)。因此波斯特认为"科学很可能最终收敛到一个唯一的真理"，并且认为"科学的进步看来好像是线性的"(同上)。 [356]

相反观点的提倡者出现在20世纪50年代后期，他们反对把"错误的连续性附加"到科学史上的思想。比如，汉森(N. R. Hanson，1958)提出，科学中的概念革命类似于格式塔转换，其中相关事实是以新方式看待的。图尔明(Stephen Toulmin，1961)也指出，巨大的概念改变经常伴随着包括这个概念在内的理论被另一个理论所代替。在他们看来，一个理论被另一个理论代替经常是革命性的颠覆。

这种激进观点的最著名倡导者是库恩和费耶阿本德(Paul Feyerabend)。他们的立场可以用不可通约性论题(incommensurability thesis)来刻画。根据这个论题，相同领域之内前后相继且相互竞争的那些理论说的是不同的理论语言。因此这些理论既不能严格比较也不能彼此翻译。不同理论的语言正是我们可以设想的不同世界的语言副本。我们能通过格式塔转换实现从一个世界到另一个世界，而不是通过任何连续过程。[1]在这点上，库恩和费耶阿本德后继有人。比如，劳丹(Larry Laudan，1981)否认本体论进步的可能性，并宣称"改变本体论或概念框架使得我们不可能抓住先前理论预设的核

心理论定律和方法”，更不用提早期本体论的基元。关于现代物理学史，皮克林(Andrew Pickering，1984)断言，1974 年之后的新物理学(大致对应于本书中的规范场纲领)和旧物理学(大致对应于量子场纲领，但也包括 S 矩阵理论)之间在最根本的意义上是不可通约的：它们属于不同的世界并且没有同一组现象。

“不可通约的”这个词本身在不同层次上指出了各种问题：主题(topic)、意义(meaning)、指称(reference)和本体论(ontology)。但是在每种情形下最终都同样宣称，在科学史上只能有“工具的进步”，这不是通常表述的越来越正确的理论命题的进步。这种论断中的相对主义和反实在论含义是很直接的。如果我们随着每一个范式转换(库恩用语)就会看到不同的世界，那么理论陈述的真值就是理论的界限，而且发现关于世界的“深层真理”这个目标在真理符合论的意义上应该是要被抛弃的东西。而且，也不会有相应于世界的真正理论命题的增加，甚至也没有朝向真理的进步。

[357] 第二种观点的温和说法是由塞拉斯、赫西和劳丹提出的。局部定律的语言不可通约性，在某些意义观的基础上遭到反对(见 12.2 节)。科学中的进步在如下论证的基础上得到承认。第一，他们争辩说，作为知识的科学主要不是宏伟的理论，而是低层定律和具体描述的全体集合。第二，这些定律和描述通过理论革命还是得到保留，因为科学有从经验学习的功能。第三，科学理论中的变化或革命通常始于概念问题而不是经验支持的问题，因此与许多低层定律不直接相关。[2]科学因此在特殊真理和局部规则性的积累上显示出进步。但是这种“工具主义进步”不同于一般定律和理论本体论的进步；因为按照赫西(1974，1985)的说法，科学中局部的或个别的成就并不需要有严格为真的普适定律和理论本体。

因此，在局部进步和本体论进步之间有重大区别。局部进步得到承认，而本体论进步遭到否定。赫西声称："没有理由假定这些理论本体显示出稳定性或收敛性，或者甚至在任何给定时刻都将是毫无争议的'最好'本体。"(1985)这也符合库恩关于理论的初始立场："我看到(在亚里士多德、牛顿和爱因斯坦体系中)本体论发展没有始终如一的方向。"(1970)

反对理论中存在本体论进步，这种观点所依据的论据是很强有力的。第一，迪昂—奎因命题，即证据对理论的非充分决定性，似乎使得真正的理论本体这一概念自身变得可疑。第二，历史记录似乎暗示着并没有理论本体的永久存在。第三，业已证明，仅当在原则上能得到一种理想地真实和普适的理论(其中理论实体"收敛"到真正实体)时，本体论进步才能被定义；然而"这将要求在未来科学中不存在概念革命，这种概念革命已经相应地被记录在直到目前为止的科学史中了"(1981；另见她1985年的论文)。

在现代科学哲学中有许多争议性问题涉及上述争论，而我不想冒昧提出本体论进步问题的解决方案。相反，我的目标不大不小：从前面几章的历史研究中总结教训，并为一种本体论进步从而为一种科学实在论和科学增长的合理性辩护。然而，为讨论本体论进步，最后需要有一个我自己的框架的总体轮廓。在我给出这个轮廓之前，对现有的适用于挑战不可通约性论题的框架作些评论可能是有帮 [358] 助的。

12.2 反对不可通约性论题的框架

为了和不可通约性论题作斗争，阿钦斯坦(Peter Achinstein, 1968)提出"语义相关属性"(semantically relevant properties)的概

念，并将其限制在感性情况中，以使物理客体能作为本体种类的成员被认识。在此概念的基础上也可证明，某些可感知属性或者说理论术语 *X* 的“语义相关”条件，都可以独立于理论被认识，并且和术语 *X* 的意义有一种特别密切的联系。因此，在此框架下，具体化一个特定本体种类的理论术语（或全称谓词），在两个理论中有同样的意义，并且在我们不知道该理论时，该理论术语的意义也是已知的。

十分相似地，在讨论描述性术语的意义变化和理论负载时，赫西(1974)按照“意向指称”(intentional reference)提出了对意义稳定性的解释。在赫西的普遍性的网络模型中，正是意向指称的概念使得人们有可能为把谓词归属到物理客体提供一个感性基础，或者说给谓词提供经验意义。由于客体之间感性上可认识的相似性以及差异，很多关于可观测属性、过程及其类比物以及经验定律的一些近似形式的理论断言，都是从一个理论到另一个理论可翻译的，从而可以既是近似稳定又是可积累的。这里，尽管一个谓词可能是理论本体论中特定种类客体的具体化，但是它还是建立在对相似性的感性认识之上。意向指称因此也是认识可观测量的经验方法，并且只有在它们产生于有谓词涉及可观测量的理论中才与不可观测量相关。

因此，在上述两个框架之内，很难找到不可观测本体论的连续性的地盘。另一个似乎能容纳这种连续性的框架是普特南的指称因果理论。[3]按照普特南，科学术语（比如水）和（水的）描述不是同义的，但是都指向具有如下属性的客体，即这些属性的产生是用包含该术语在内的适当类型的陈述的因果链联结在一起的。这个理论借助于“疑义利益原则”(principle of benefit of the doubt)，可扩大到涵盖科学中的不可观测术语（比如电子）。这个原理说的是，用来把一个[359]不可观测术语的指称具体化但不能实指的描述（比如玻尔对电子的描

述）合理地再公式化总是可被人们接受的。这种再公式化（比如量子力学中对电子的描述），使那些指向某类客体的早期描述，从稍晚理论的立场来看也存在几分相同作用（Putnam，1978）。因此我们能够谈论不可观测术语的跨理论指称，并且能根据这个指称稳定性来解释意义稳定性。和不可观测术语在不同理论中不可能有同样的指称这个基本观点相反，指称的因果理论允许我们谈论不同理论——那种可以通过不可观测术语来确定同类客体的理论（同上）。

普特南（1978）进一步论证道，如果没有先前理论是后来理论的极限情况这个意义上的收敛，那么“疑义利益”就绝不可能是合理的。说它不合理是因为像燃素这种理论术语曾不得不用于指称，而又不可能在理论发生了巨大变化后仍保留这种指称。另外，普特南争论说，相信收敛将导致尽可能经常地保留早先理论的机制的方法论，而这将限制候选理论的种类，从而增加成功的机会。

这种限制的方法论含义放在 12.5 节和 12.6 节讨论。这里我只是要对普特南的指称因果理论作点评论，这是与我所采用的容纳本体论进步的框架相关的。普特南（1978）正确地看到，与特定理论术语的指称相关的，不仅有假说性的指称对象特有的属性（电荷、质量等），而且还有所解释的效应以及假说性的指称对象所起的作用。即使我们可以把理论实体特有的属性通过与可观测客体的性质相类比而解释成感性的，在普特南看来，其他由“效应”和“作用”所暗示的性质，并不直接是感性的。在我看来，这些属性与我所称的“结构属性”（structural properties）或实体的性质和关系的属性有关联。这些结构属性，虽然并不直接是感性的，但仍然是特定理论术语的指称的关键，并因此在决定理论术语的意义时起到决定性作用。

12.3 本体论的结构属性

在勾勒我自己所描述的经由概念革命的本体论进步的框架之前，首先应该回答一个问题：如果科学主要是经验定律的全体集合，为什么在科学中还需要有本体论呢？ 答案是，绝大多数职业科学家相信经验定律只有“局部”有效性，而本体论赋予科学以统一能力。 本体论作为世界基本结构的模型，是经验定律的基础的一般机制的载体。 由于在它的基础之上，可以建立一门科学的统一概念框架，所以本体论在理论上比个别经验定律更为基本得多。 [360]

本体论的理论术语不仅在智力虚构中有用，而且在某种意义上也是有所指的。 石里克(1918)、罗素(1927)、卡尔纳普(Rudolf Carnap，1929)以及马克斯韦尔(Grover Maxwell，1971)都注意到，而我也同意，本体论的陈述所指向的是根本性的特殊实体，以及主要通过这些实体的结构特征指称的内在属性。[4]然而，靠这些陈述，我们还是无法确切地知道这些实体是什么，以及这些内在属性和结构属性是什么。 不过我们至少能知道实体的存在及其内在属性和结构属性的存在。 这里的内在属性(或关系)是个体的一阶属性(或关系)，并且是一元(或多元)谓词符号的直接指称；而结构属性是性质和关系的属性：它们是高阶逻辑类型的指称。

在与有些如此这般的特殊实体相关的属性中间，有所谓的基本属性(essential properties)。 这些属性在理论中定义，并且它们的描述在从一个理论到另一个理论的转换时可能会发生改变。 一旦人们发现以前的基本属性可以被新的基本属性解释，那它就不是基本的了。 这个理论依赖性的一个结果是，只要科学继续发展，一个特定实体在本质上是什么将永远不会得到最终解决。 因此实体意义上的理论本体论不能被看作真理性回答，而只能作为物理实在的模型，建立在由

我们根据所知而提出的类比的基础上，并最终由观测决定。

因此，这样一个模型实际上提供给我们的不是根本性实体在字面上的正确描述，而是通过类比由假说性实体承载的可观测结构关系的断言。事实上，一种更强的说法认为，系统的任何本体论特征在本质上总是唯一地结构性的。那就是说，本体论的作用主要体现为所建立的基本实体的结构关系。因此，结构属性和关系是作为本体论的一部分，而且，它们是进入科学研究者那里的本体论的唯一一部分，这种进入是通过将结构断言和假说性实体联系起来的因果链。[361] 虽然结构断言在理论发生变化和新结构属性被发现时也要加以修正，但像可观测量这些属性既是近似稳定的也是渐增的，因为它们借助于其可认识的同一性从一个理论到另一个理论之间是可翻译的。

一些关于结构属性的例子是我们这一章讨论的关键，其中包括外部对称性和内部对称性、可几何化性，以及可量子化性。满足物理对象的定律的外部对称性（如洛伦兹对称性）明显在本质上是结构性的。事实上，变换群的基本作用和关于把变换群作为客体集体结构特性描述的不变性思想，是庞加莱从19世纪（1895，1902）就意识到的和倡导的，并且随后融入像爱因斯坦、狄拉克、维格纳、杨振宁和盖尔曼等数理学家的集体意识之中。[5]内部对称性（比如同位旋对称性）是通过诺特定理和物理客体（比如电荷、同位旋荷等守恒量）的内在属性联系在一起的，但没有全部联系起来。然而，作为抽象内部空间中的对称性，它们也是一些高阶属性。可几何化性是和时空流形或其扩展的结构特性同构的结构属性。可量子化性是连续充盈的结构属性，这是和离散可以从连续中创造或湮灭于连续中的机制联系在一起的。

因此，有关本体论的陈述包括指称（在网络意义上定义的）可观测

量的术语和逻辑术语。也就是说，本体论的潜在指称是由其可观测结构属性确定的。比如，粒子本体论的指称部分地是由诸如“物理客体具有可隔离的组分”以及“这些组分具有某种理论上可表征的自主性”这些结构属性确定的。场本体论的指称是通过诸如“实体的不同部分之间的可叠加性”和“把实体个体化的不可能性”之类的结构属性而确定的。

在理论本体论的结构属性之间是诸如几何化原理、量子化原理和规范不变性原理之类的指导原理。这些原理产生并统一经验定律（关系）；因此，它们的连续性能用经验定律的连续性辩护。下面我将把我的讨论限制在这些指导原理范围之内。

在科学理论中，已经发现的结构属性假定是由理论实体承载的。[362] 随着结构属性的积累，旧的理论实体不可避免地会被新的理论实体和整个理论的本体论特性的改变所取代。然而，只要涉及旧的结构属性，本体论的变化仅仅意味着它们在结构属性整体中的作用和位置的改变。因此，理论实体的替换应被适当地看作理论本体论的类比扩展，这是由结构属性的积累引起的，而不是革命性的推翻所引起的。

12.4 经由本体论综合的概念革命

我在上一节所作的评论已经提供了一个可以用于讨论本体论进步的框架。关于本体论进步，我指的是通过关于世界的一系列理论所展示出的结构关系的积累和扩展。理论的统一能力的根源严格说依赖于如下事实：实体意义上的理论本体是一种隐喻，并且具有可以称为隐喻扩展的能力，而且在结构属性的意义上，本体在本质上是稳定的和累积的。[6]

乍看之下，几何纲领和量子场纲领在描述基本相互作用时是如此

不同，以致可以把它们当成不可通约性论题的典型情况。 前者，代表传递相互作用的物理实体的表象在本质上是几何的，在时空中是连续的；而后者，相应的表象是离散量子。 前者，广义坐标协变性是指导原理；而后者，对量子行为不作连续的时空描述也是可能的，并且特定的坐标系不由分说就设定了。 在这两个纲领之间的对比是如此分明，以致没有人能否认的确发生了概念革命。 那么，谁敢说旧本体论的几何结构能被结合到新本体论的量子之中呢？ 我们在什么地方能看到本体论发展的始终如一的方向呢？

的确有概念革命，不过其本体论转换的含义正是引起争论的主题。 如果我们利用过后方知的优势，在更深的层次上探索这两种纲领的本体论基础和潜在可能性，那么整个图景就将与库恩和费耶阿本德提供的十分不同。

前面几章中的历史研究能够概述如下。 20 世纪场论的所有方案都源自经典电动力学(CED)。 经典电动力学是实体电磁场理论，并且具有洛伦兹群作为其对称群(见 3.3 节)。[7]

[363] 引力场论是直接继承经典电动力学的：爱因斯坦把洛伦兹不变性(整体外部对称性)推广到广义协变性(局域对称性)，并通过等效原理引入引力场，两者结合发展了他的广义相对论(GTR)(见 3.4 节和4.2节)。 广义相对论开创了几何纲领，其中实体场与时空的几何结构不可分离地联系在一起，相互作用通过时空传递。 按照几何纲领的弱版本，其基本本体也是实体场：时空及其几何结构自身是没有存在性的，而只是作为场的一种结构上的质(见 5.1 节和 5.2 节)。 这里，几何纲领和经典电动力学之间的本体连续性是明显的。

量子电动力学(QED)是经典电动力学的另一个直接继承者。 当经典电磁场被量子化电磁场所代替，并且通过与玻色子场的类比，实体

费米子场被引入，量子电动力学为物理学界所使用(见7.3节)。 量子电动力学开创了量子场纲领，其中场以传递相互作用的离散量子的形式显现自身，在这个纲领中基本本体也是某种实体场(见7.3节)。

量子场纲领中的实体场(量子场)，其结构属性根本不同于经典电动力学中的实体场(即经典场)，并且这个本体论差异使它们属于不同的范式，这一点是正确的。 不过，量子场和经典场还是共享了“不同成分的叠加性”和“个体化的不可能性”之类的硬核式结构属性。一般说来，在任何两个具有似乎不同的本体的理论所共有的硬核式结构属性的基础上，我们总能建立这两个理论之间的本体对应[8]，从而在结构属性的意义上，使本体在这两个可区分理论中的指称连续。尤其是，借助于玻尔的对应原理，我们能在经典电动力学和量子电动力学之间的本体变化中找到一种连续性，那就是，在量子场和经典场两者结构属性方面之间的一种指称连续性。

规范理论直接继承了量子电动力学。 杨振宁和米尔斯所做的仅仅是用局域SU(2)同位旋对称代替了量子电动力学的局域U(1)相位对称，前者是假定用来描述强相互作用的。 杨—米尔斯理论开创了规范场纲领，其中基本相互作用的形式是由规范不变性的需要固定的。 在规范场纲领和量子场纲领之间的本体连续性说法没有受到挑战，因为规范场跟“物质场”一样，也是被量子化的。 但规范场纲领的本体和几何纲领的本体之间的关系又是什么样的呢?

[364] 像量子电动力学和经典电动力学之间的情况一样，这两个纲领之间的本体连续性在于量子化规范场和经典几何场在它们的结构属性中的指称连续性。 不仅几何纲领中的理论，比如广义相对论或者它的推广和变体，(至少在原则上)能被写成量子化形式[这很类似于规范场纲领中理论的数学结构(见9.3节、11.2节和11.3节)]，而且人们

也能给规范场纲领中的理论以几何解释(见11.3节)。按这种方式能证明，这两种纲领的本体有共同的结构属性甚或基本特征。规范场纲领和几何纲领之间的本体连续性以这样一种特别直接的方式显现在现代卡卢察—克莱因理论中，以致我们能毫不犹豫地宣称规范场纲领是几何纲领的直接继承者(见11.3节)。

注意到下面这一点是有趣的，即量子场纲领的本体和几何纲领的本体之间的密切联系，只有在规范场纲领达到作为这两者的综合的阶段(见11.3节)之后才变得清楚。这个事实告诉我们，综合的概念对于认识理论本体在经历概念革命时的指称连续性是有帮助的。

这里科学思想的“综合”并不意味着以前的思想或原理在一个“综合的混杂物”中的联合，而是指一种更高级的选择性联合，此时预先假定了对以前的科学思想(概念、原理、机制、假说等等)的变换。以前的每一个思想中的一些有用因素都被保留，而其余的被舍弃。只有当所保留的思想被变换并且成为本质上新的不同思想，综合才成为可能。比如，牛顿综合了开普勒的惯性概念(作为物质的一种性质，只有在使它们运动的力停止作用时物体才静止)和笛卡儿的惯性运动概念(作为自然律而不是运动定律，也就是任何事物总是保持相同状态)，并且形成他自己的惯性概念，即作为物质属性，使物质保持它们所在的任何状态，不管是静止状态还是匀速直线运动状态。在这种综合中，以前的两个概念全部得到了改造。

我们可以把综合的概念扩展到关于本体论的讨论，并发现本体论综合也以改造为先决条件。事实上，几何纲领的诞生是综合了广义协变性思想和等效原理的结果，而广义协变性思想源自于但又不同于洛伦兹不变性的思想；量子场纲领的诞生是综合了经典电动力学和量子原理的结果，而场量子化的思想源自于但在本体论上又不同于原子

[365] 运动的量子化思想(见 6.1 节、6.2 节、7.2 节、7.3 节)；规范场纲领的诞生是综合了量子场纲领和规范原理的结果，而规范原理源自于但又不同于最小电磁耦合(见 9.2 节)。在所有这三种情况下，以前的原理都转换成新的形式，并且只有在那时才对本体论综合有用。

而且，作为一般特征，本体论综合通常把实体(作为原初实际存在物)变成副现象(或导出的实际存在物)，因此伴随着基本本体的变化。比如，在几何纲领中，牛顿引力势被看作度规场(的一部分)的显示；在量子场纲领中，经典场被看作量子场的一个副现象；在规范场纲领的标准模型中，量子电磁场被看作量子化规范场的一个副现象(见 11.1 节)。因此，正如这一节所概述的，在这三个纲领之间清楚的线索暗示着本体论综合是完成概念革命的可能道路之一。综合的结果即是新研究纲领的诞生，这是建立在新的基本本体的基础上。关于这个观点，正如内格尔(1961)和波斯特(1971)所认为的，革命前纲领的本体跟直接结合到革命后纲领的本体是很不一样的。

另一方面，已发现的关于世界的结构关系(比如，外部对称性和内部对称性、几何化、量子化等)，通过概念革命，在极限意义上体现为对旧本体论的坚持。通过变换和综合完成的概念革命绝不是绝对的否定，而是可看作黑格尔意义上的“扬弃”(Aufhebung)，即有保留地克服、改变的意思。因此，科学不仅是以经验定律的增加的形式显示进步，而且甚至更有意义地是以概念革命的形式显示进步。科学的目的在于富有成效的隐喻和更详细的世界结构。在概念革命后我们就更接近这个目标，因为随着每一次革命，经验定律就会被一个比旧本体论更好的新本体论更好地统一起来。

正是在这种保留和积累世界结构的意义上，我们声称本体论的发展是指向世界的真正结构。注意，这里的表述“世界的真正结构”

应该理解成“一个尽管总是可扩大的、但却是确定的研究领域的真正结构”。由于没有理由假设世界的结构性质只能被分成有限层次，而又没有理论能够抓住无穷数目的结构性质，所以关于为我们提供整个世界真正结构的终极正确的本体概念，显然是无意义的。相应地，在上一段中的用语“更好”和“更接近”只有旧理论和新理论之间相比较的含义，而没有绝对的含义。 [366]

12.5 概念革命与科学实在论

在这一节中我将转到如何把在上一节中得出的启示建立在科学实在论基础上的问题。按照已经接受的观点[9]，科学实在论假定：(i)不断进步而可接受的理论所预设的实体及其结构属性确实存在，虽然不能肯定理论所给出的对实体的每一特定描述都为真；(ii)理论的目标在于得到与世界相符合的真命题；以及(iii)至少成熟科学[10]的历史显示出一种趋向于真理的进步；在相应的意义上，考虑所研究的领域，也显示出一种既在“低层次”的定律方面，甚或更重要的，也在支撑“低层次”定律的本体方面(在局部但可扩展的结构关系的意义上)的进步。第三点和本书的历史研究最为相关。

科学实在论要求本体在经过概念革命后有连续性。否则，一个范式中的理论本体就没有在下一个范式中要求存在的任何权利，因而实在论就完全错了。反实在论对第三点的反驳涉及本体论。对他们而言，概念革命的存在意味着理论科学在历史上是根本不连续的，并且给出了对上述定义的实在论一个强有力的反对论证。要是成功的科学理论受到预设了根本不同的本体的理论的拒斥，其中只有以前认识了的低层次定律得到保留，那么如此定义的科学实在论似乎不是一个站得住脚的立场。

普特南(1978)在著名的“元归纳”(meta-induction)讨论中曾对此情况作了概括。如果概念革命使得科学史是绝对不连续的，那么“最终下面讲的元归纳就会令人信服：正如没有一个术语按照50年前的科学来指称一样，同样也没有一个术语在将来会按现在的意义来指称”(同上)，或者说科学实在论是一门错误的学科。

对概念革命的本质进行解释对科学实在论是如此关键，以致实在论者不得不寻找论证来阻止灾难性的“元归纳”。对实在论的论证有一种是由普特南在其指称因果理论中提供的。逻辑上，指称因果理论借助于“疑义利益原则”，有可能把科学革命描述成包括了理论术语的指称连续性。[367] 显然，借助这样一种指称连续性，灾难性的“元归纳”能被成功地阻止，并且科学增长的单线(unilinear)观，也就是内格尔和波斯特所采取的特殊收敛实在论，能或多或少得到了辩护。

但是这种对反实在论论证的反驳，既过于简单又太抽象。说它过于简单，是因为科学进步的结构比连续性的单线性形式所构想的更丰富，虽然指称连续性的概念能够用作概念革命的进步性的实在论概念的基础。说它太抽象，是因为所涉及的逻辑证明，既没有解释概念革命的出现也没有解释理论的发展，特别是当涉及新的不可观测实体的假定时。因此，这种反驳不能令人信服地用实际科学发展的历史分析来支持实在论。

另外一个实在论的论证是由普遍性网络模型提供的(Hesse, 1974)。逻辑上，普遍性网络模型使得人们有可能把科学革命描述成包含理论术语的相似连续性。显然，有了这样一个相似连续性思想，再加上认识到理论术语的指称必须从可观测量导出并且得到意向性理解这一点，就可以成功阻止灾难性的“元归纳”。在这种论证

的基础上，“工具”进步的弱实在论得到了辩护(同上)。

我对这个问题的态度是赞成通过引入本体论综合的概念，超越特殊收敛实在论和弱实在论。这个概念是直接从对20世纪场论的历史分析中得到的，因此不能被指责为与实际科学实践无关。而且，它也能用来解释经由概念革命的进步的复杂形式。科学的增长和进步不必把它们自身表述成连续性和积累的单线性形式。本体论综合的概念作为已发现的世界结构的连续性和积累的辩证形式，在解释概念革命的机制和科学进步的模式方面更有效。事实上，科学中的连续、累积和统一很少是以直线发展的形式实现的，相反，常常是通过经由变换的综合的形式实现的，正如上一节所概括的。

总之，我们可以说本体论经由变换的综合这一概念作为实现概念革命的一种可能方式，代表了科学增长中的变化和保守之间的调解。因此，它适合在科学史中通过表观上的不连续表现抓住根本的连续性。在这一概念的基础之上，可以发展出一种较强的科学实在论论证。

12.6 概念革命与科学合理性

[368]

科学合理性(scientific rationality)这个概念[11]与真理概念是紧密联系在一起的，只有把科学的核心目标指向关于经验定律和世界结构属性的越来越正确的论断，这种科学进步的特征才允许我们把科学作为理性活动的代表。在这个意义上，实在论可以被视为基础，虽然对科学合理性而言不是必然的唯一基础。另一方面，科学推理在决定哪个理论获得了对世界的正确描述方面起了关键作用。因此，这对通过其真理性断言来评价理论以及在相互竞争的理论中作出最终选择，都是必要的。[12]

对科学合理性的威胁来自革命性的范式转换的观点。有了这个观点，科学革命就可以和信仰的改变或者格式塔转换的现象相比较。事实上，这种比较成为不可通约性论题和反对真理概念以及本体论进步概念的主要论证方式。在这个意义上，科学革命的思想使科学合理性成为疑问，并且产生科学合理性的危机。这十分令人担忧，原因之一是这最终会使合理性理论的评价和理论选择变得不可能。然而科学在理论发展中延续。这对基础物理学尤其是真的，其中重大进展通常采用包括本体变化在内的概念革命的形式，但是又没有明显的非理性。因此建立在不可通约性思想基础上的科学哲学将和实际的科学活动不相关。

因此，为了让科学合理性抵御这种威胁，在对理论所研究的领域的描述越来越正确的意义上，人们不得不认为概念革命是进步的。正好处在这个争论的关节点上的是，我发现把本体论综合的概念作为一种实现概念革命的方法，对争论起到关键作用。这里的进步是按照理论结构(本体、机制、原理等等)的结构关系的扩展方式定义的，它们产生了连续性，但不是特殊收敛的。

把合理性观点建立在本体论综合概念的基础上，使得由科学进步的综合观提出的基础研究未来方向，不同于由不可通约性观所提出的方向。因为对某些持有后一种观点的信徒来说，科学研究的方向主要是由外部因素，比如社会因素或心理因素决定的，而跟智力因素关[369]系不大。对持有其他观点的人而言，智力考虑更重要，因为科学演化有爆发式特征，它起到一种本质上不可预测的作用。相反，综合观则要求先前理论的内在方法必须尽可能保留。

这种立场的另一个重大意义是认为科学增长不是单线性的，而是辩证的。新的实践总是产生新的数据，因为世界的方方面面在数量

上其实是无穷的。由于新数据而出现的新观念，常常伴随现存观念的变化，因此总是需要一次新的综合以及一次新的概念革命。因此，在此意义上科学进步和未来的概念革命并非不相容。革命是永久性的。收敛于一个固定真理与进步的综合观才是不相容的。

另外还有一个差异存在于综合观和单线观之间。按照后者，现存的成功理论一定是未来发展的模型。比如，爱因斯坦的广义相对论曾经被指望能导致对电磁和物质的理解。众所周知，这个期望破灭了。事实上，科学史中对先前模型的放弃已经反复发生。流行于17、18世纪的“机械论哲学”后来被抛弃了；流行于19世纪末的“电磁世界观”也在20世纪被抛弃了。在20世纪20年代，在狄拉克发表了关于电子的相对论性波动方程之后不久，玻恩声称“就我们所知的物理学将在6个月内完成”。[13]所有这些希望都证明失败了。不幸的是，同样的事情近年来又重复发生。这次充当模型的成功理论是规范场纲领的标准模型。[14]规范场纲领的命运又会怎样呢？事实上，规范场纲领远非完备的理论框架，即使单单拿希格斯机制和汤川耦合来说，都不能用规范原理来解释。[15]

相反，由综合观提出的关于未来研究的建议在总体上完全不同。它劝导科学家对所有可能性都要保持开放的心态，因为超出现存概念框架的新综合总是可能的。按照这种精神，场论的未来发展不能完全排除在规范场纲领研究之外，并且应试图把安德森—希格斯机制和汤川耦合结合到此纲领中。很有可能它们会被发展起来的思想和技巧的某种使用所激发，比如，在S矩阵理论中，其诸如本体论和力的本质之类的基本思想，根本不同于规范场纲领中的情况。

总之，科学合理性在于意图得到物理世界的真正结构的近似知识；而在科学增长的综合观看来，实现这种合理性的一条道路就是通 [370]

过概念革命。

注释

1. Kuhn(1962)和 Feyeraband(1962)。

2. 例如，见 Sellars(1961，1963，1965)；Hesse(1974，1981，1985)；Laudan(1981)。

3. Putnam(1975，1978，1981)；另见 Kripke(1972)。

4. 关于本体论的结构主义观的进一步讨论，见 Chihara(1973)，Resnik(1981，1982)，Giedymin(1982)，以及 Demopoulos and Friedman(1985)。

5. 例如，见 Einstein(1949)，Dirac(1977)，Wigner(1949，1964a，b)，Yang(1980)，和 Gell-Mann(1987)。

6. 另见 McMullin(1982，1984)和 Boyd(1979，1984)。

7. 这里的理论对称群被理解成理论的标准表述的协变群，这只是理论自身的动力学客体的微分方程系统，见 Friedman(1983)。

8. 当然，这不是没有争议的。关于理论变化，存在大量从不同立场处理这个论题的文献。例如，见 Feyerabend(1962)，Sneed(1971，1979)，Post(1971)，Popper(1974)，Kuhn(1976)，Stegmüller(1976，1979)，Krajewski(1977)，Spector(1978)，Yoshida(1977，1981)，以及 Moulines(1984)。关于此论题近来的贡献，见巴尔策、皮尔斯和施密特(Balzer，Pearce and Schmidt，1984)的研究。

9. 见 Putnam(1975)和 Hacking(1983)。

10. 成熟科学包含一个或一个以上的成熟理论，而成熟理论以一个连贯的数学结构、有效领域(经验支持)以及横向(根据它和其他不同分支的理论之间的关系)和纵向(根据它和其他成熟理论之间的关系)的连贯性为特征。所有下面涉及科学实在论的讨论，只是就成熟科学而言才有意义。关于成熟科学的更多讨论，见 Rohrlich and Hardin(1983)。

11. 这一节中关于合理性的讨论限于科学合理性的概念，不包括合理性的其他形式，比如实践上或美学上的考虑。应该指出的是，库恩，这位在此节中被批评的人，已经在近年来(1990，1991，1993)发展了科学事业中实践合理性(practical rationality)这样一个意义深远的概念。

12. 关于理论选择的问题，我假定了一种强内在主义立场，根据这个立场，最终的经验上和逻辑上的考虑在作出选择时起了关键作用。然而，它太复杂且不太相关，因而不能在简短的哲学讨论中抵御社会建构论者的攻击。不过，见 Cao(1993)。

13. 转引自 Steven Hawking(1980)。

14. 例如，见 Hawking(1980)和 Glashow(1980)。

15. 与规范场纲领有关的问题的进一步讨论可在 11.4 节找到。

附　　录

附录 1. 内蕴几何学、局域几何学和动力几何学的兴起 [371]

A1.1　内蕴几何学

在高斯致力于微分几何的研究(1827)之前，平面一直被作为三维欧几里得空间中的图形来研究。但是高斯表明，关于平面的几何学能通过专注于平面本身得以研究。平面 S 是具有两个自由度的点的集合，因此 S 上的任意一点 r 能用两个参数 u_1 和 u_2 来表示。我们能获得表达式：$dr = (\partial r/\partial u^1)du^1 + (\partial r/\partial u^2)du^2 = r_i du^i (r_i \equiv \partial r/\partial u^i$，关于 $i = 1, 2$ 时的求和缩写法)，并且 $ds^2 = dr^2 = r_i r_j du^i du^j = g_{ij} du^i du^j$。高斯得出观测结论：平面的属性，如弧长元、平面上两条曲线之间的夹角，以及通常所谓的平面的高斯曲率，仅仅取决于 g_{ij}，而这有许多推论。如果我们引入坐标 u^1 和 u^2——这来源于三维空间中平面的参数表示 $x = x(u^1, u^2)$，$y = y(u^1, u^2)$，$z = z(u^1, u^2)$——并且运用由此确定的 g_{ij}，我们就获得这个平面的欧几里得性质。但是，我们能从这一平面出发，引入

两组参数曲线 u^1 和 u^2，并用 g_{ij} 作为 u^1 和 u^2 的函数而获得 ds^2 的表达式。因此，这一平面有一个由 g_{ij} 确定的几何学。

这一几何学是内蕴于平面的，并且与周围的空间没有关系。这表明这个平面本身能被看作一个空间。如果这个平面本身被看作一个空间，那么它拥有什么类型的几何学呢？如果我们认为那个平面上的“直线”是测地线(平面上两点间的最短连线)，那么此几何学可能是非欧几里得的。因此，高斯的工作所隐含的是，至少在本身被看作空间的平面上有非欧几何。

受高斯关于欧几里得空间中平面的内蕴几何学的引导，黎曼为一个种类更为宽泛的空间发展了一种内蕴几何学(intrinsic geometry)(1854)。尽管三维几何学显然是重要的几何学，但是黎曼更喜欢处理 n 维几何学。他把 n 维空间当作一个流形来讨论。在一个 n 维流形中的点由赋予 n 个变元参数 x_1，x_2，…，x_n 的特定数值，和构成 n 维流形本身的所有这些可能的点的总数表示。同高斯的平面内蕴几何学一样，黎曼流形的几何学性质是用流形自身可确定的量来定义的，并且没有必要把流形看作位于某种更高维的流形之中。[372]

A1.2 局域几何学

根据黎曼，

> 我们对于现象的因果关系的了解，本质上依赖于我们对于无穷小现象理解的精确性。近世纪以来，在力学知识上的进展几乎完全依赖于构造的精确性，这种精确性已通过无穷小积分运算的发明而变得可能。
>
> (Riemann，1854)

因此，“关于在无穷小空间中的测量关系的问题”是最重要的（同上）。

与欧几里得的有限几何学相比，黎曼几何作为本质上一种无限近点几何学，符合莱布尼茨关于连续性原理的思想，根据连续性原理，没有一个相互作用定律能用超距作用来表述。因此，黎曼几何能与法拉第关于电磁现象的场概念相比较，或与黎曼自己关于电磁、引力和光的以太场论相比较（参阅 2.3 节）。外尔曾把这种情形描述为：“从无穷小部分的行为获得外部世界的知识的原理，是无穷小物理学中也是黎曼几何中知识理论的主要推动力。”(1918a)

A1.3 动力几何学

黎曼指出，如果物体依赖于位置，“我们就不能从极大物体的度规关系到极小物体的度规关系中得出结论”。在这种情形中，

> 空间的度规测定所建基于其上的经验概念——固体的概念和光线的概念——对于无穷小量似乎不再有效。因此，我们可以相当自由地假定，无穷小空间中的度规关系不遵循（欧几里得）几何的假说；如果我们能由此获得关于现象的一个更为简单的说明，那么我们实际上就应该这样假定。
>
> (Riemann, 1854)

这就表明，物理空间的几何学，作为一种特殊的流形，不能只是从关于流形的纯粹的几何概念中导出。区分物理空间和其他三重延展流形的性质，将只得从经验中获得，即通过引入测量仪器，或通过拥有一种关于以太力的理论等等而获得。黎曼继续说道：

关于在无穷小量中的几何假说的有效性问题，是与空间的度规关系的基础问题缠结在一起……在一个连续流形中……我们必须在这一流形之外，在作用于其上的束缚力之中，寻求这一流形的度规关系的基础……这把我们引入另一门科学的领域，即物理学的领域。

（同上）

[373] 这段话清楚地表明了黎曼拒绝如下这样一种观念：关于空间的度规结构是确定不变的，而且它作为物理现象的一个背景，天生独立于物理现象。相反，他断言，空间本身只是一个缺乏所有形式的三维流形：只有通过给其填充物质内容并确定其度规关系，它才获得了一个确定的形式。在此，物质内容由他的以太理论描述(见 2.3 节)。考虑到世界上物质的配置会发生变化这一事实，因此度规的基础形式也会随时间改变。黎曼这种度规依赖于物理数据的预期，后来为避免绝对空间的概念提供了辩护，绝对空间的度规不依赖于物理力。例如，60 多年后，爱因斯坦把这里提到的黎曼关于几何的经验概念，看作对他的广义相对论的一个重要辩护(见 4.4 节)。

黎曼把物质和空间联系起来以确定物理空间什么是真实的这个思想，进一步为克利福德所发展。对于黎曼而言，物质是空间结构的动力因；在克利福德看来，物质及其运动显现了空间的变化曲率。克利福德这样表述："从一点到另一点可能会出现曲率的轻微变化，它们自已随时间而变……我们甚至可以进入到给它们的空间曲率变化赋值的层面，'在我们称作物质运动的现象中，真的发生了什么'。"(在 Newman(ed.)，1946，p.202) 1876 年，克利福德发表了一篇论文"论物质的空间理论"(On the space-theory of matter)，他

在其中写道：

> 我认为，事实上，(1)空间的小部分实际上与一座平均来说是平坦的平面上的小山丘性质相似；也就是说，在这些空间和平面上，几何学的常规定理不再有效。(2)这一弯曲的或扭曲的性质，仿照波的方式，连续地从空间的一个部分传递到另一个部分。(3)空间曲率的这一变化，在我们称作物质运动的现象中确实发生，不论这些物质是可称重的还是以太的。(4)在物理世界中，除了这种(可能)服从连续性定律的变化外，没有其他任何事情发生。
>
> (1876)

显然，所有这些想法都受到了黎曼以太场论的强烈影响，尽管黎曼的以太为克利福德的空间所重新命名。

A1.4 不变量

黎曼在他1861年的论文中，阐述了度规 $ds^2 = g_{ij}dx^i dx^j$ 何时能通过方程 $x_i = x_i(y_1, \cdots, y_n)$ 转化为给定度规 $ds'^2 = h_{ij}dy^i dy^j$ 的一般问题。对于这一问题的理解是 ds 等同于 ds'，以使除了坐标的选择外，这两种空间的几何学相同(1861b)。1869 年，克里斯托费尔(Elwin Christoffel)在他的两篇论文中重新思考和阐述了这个主题。在这两篇论文中，他引入了克里斯托费尔符号。克里斯托费尔表明，对于 μ 阶微分形式 $G_\mu = \sum(i_1, \cdots, i_\mu)\partial_1 x_{i1}, \cdots, \partial_\mu x_{i\mu}$ 来说，下面的关系式 $(\alpha_1, \cdots, \alpha_\mu) = \sum(i_1, \cdots, i_\mu)\partial x_{i1}/\partial y_{\alpha 1}, \cdots, \partial x_{i\mu}/\partial y_{\alpha\mu}$ 对于 $G_\mu = G'_\mu \equiv \sum(\alpha_1, \cdots, \alpha_\mu)\partial_1 y_{\alpha 1}\cdots\partial_\mu y_{\alpha\mu}$ 来说是充要的，其中 $(\mu+1)$ 指标符号能通过里奇和列维-齐维塔后来叫做“协 [374]

变微分”的方法，从一个按照 g_{rs} 定义的 μ 指标符号 $(i_1, \cdots, i_\mu)$ 中导出。

上述研究隐含着，对于完全相同的流形来说，可获得不同的坐标表象。然而，流形的几何性质必须独立于用来表象它的特定坐标系。这些几何性质用不变量解析地表示。黎曼几何中关心的不变量包括基本的 ds^2 二次项形式。据此，克里斯托费尔推导出高阶微分形式，他的 G_4 和 G_μ 也是不变量。而且，他还表明了如何能从 G_μ 导出另一个不变量 $G_{\mu+1}$。

不变量的概念得到了费利克斯·克莱因(1872)的拓展。对于任何 S 集合及其变换群 G 来说，如果对于每个 x 有 $x \in S$ 且 $f \in G$，无论何时 x 有性质 Q，$f(x)$ 就有 Q，那么我们说变换群 G 保持 Q。同样我们可以说，G 保持一个在 S^n 上定义的关系或函数。由 G 保持的任何性质、关系等，都被认为是 G 不变的。费利克斯·克莱因用这些思想来定义和澄清几何学的概念。令 S 是一个 n 维流形，G 是关于 S 的变换群。通过把 G 加到 S，费利克斯·克莱因定义了关于 S 的一种几何学，它包含在 G 不变的理论中。因此，一种几何学不由定义它的流形的元素的特殊性质确定，而是由定义它的变换群的结构确定。

不是所有的几何学都能被并入费利克斯·克莱因的框架之中。黎曼的流形几何学就不适合于这一框架。如果 S 是一个非常数曲率的流形，弧长可能碰巧就不是由 S 的变换群所保持，而是由单位矩阵单独组成的平凡群所保持。但是，这个平凡群不描述任何东西，更不用说黎曼关于流形 S 的几何学了。

A1.5 张量运算

求微分不变量的一种新方法由里奇开创。里奇受沿着克里斯托

费尔的方向工作的比安基(Luigi Bianchi)的影响。里奇和他的学生列维-齐维塔一起设计出了新方法，并给出了这个主题一个易于理解的概念，他们称之为绝对微分运算(1901)。

里奇的思想是：不是把注意力集中在不变微分形式 $G_\mu = \sum(i_1, \cdots, i_\mu)\partial_1 x_{i1}\cdots\partial_\mu x_{i\mu}$ 上，而是应该充分地且更敏捷地处理 n^μ 分量($i_1\cdots i_\mu$)的集合。他把这个集合叫做(协变或逆变)张量，只要它们依照某种(协变或逆变)规则在坐标改变下变换。在一个坐标系中一个张量拥有的物理和几何意义为这一变换所保持，以致能在另一个坐标系中再次获得。在1901年的论文中，里奇和列维-齐维塔表明物理定律如何能用张量形式表达，以使它们不依赖于坐标系。因此，正如爱因斯坦在其广义相对论的表述中所做的，张量分析能被用来表达由相应的坐标系表示的所有参考系所拥有的物理定律的数学不变性。

张量的运算包括加法、乘法、协变微分和缩并。通过缩并，里奇从黎曼—克里斯托费尔张量中获得了现在所称的里奇张量或爱因斯坦张量。分量 R_{jl} 是 $\sum_{k=1}^{n} R_{jlk}^{k}$。这个 $n = 4$ 的张量被爱因斯坦用来表达他的时空黎曼几何曲率。 [375]

附录2 同伦类与同伦群

A2.1 同伦类

令 f_0 和 f_1 是两个从一个拓扑空间 X 进入另一个拓扑空间 Y 的连续映射。如果它们是相互连续可变形的，那么就可以说它们是同伦的。也就是说，当且仅当存在映射 $F(x, t)$，$0 \leqslant t \leqslant 1$ 的一个连续变形，诸如 $F(x, 0) = f_0(x)$ 且 $F(x, 1) = f_1(x)$，函数 $F(x, t)$

才被称作同伦的，X 到 Y 的所有映射才能被分成同伦类。如果两个映射是同伦的，那它们就属于同一类。

A2.2 同伦群

在同伦类集上能定义一个群结构。最简单的情形如下。令 S^1 是一个用角 θ 参数化的单位圆，θ 和 $\theta+2\pi$ 被看成同一的。因此，从 S^1 映射到一个李群 G 的流形中的同伦类的群，被称作 G 的第一同伦群，用 $\Pi_1(G)$表示。如果与 S^1 拓扑等价的 G 是用一个幺模复数 $u=\exp(\mathrm{i}\alpha)$ 的集表示的 U(1)群，那么 $\Pi_1(\mathrm{U}(1))$(或从 S^1 到 S^1 的连续函数)的元素 $\alpha(\theta)=\exp[\mathrm{i}(N\theta+\alpha)]$，对于不同的 α 值和一个确定的整数 N 来说，形成一个同伦类。$\alpha(\theta)$可以看作一个圆到另一个圆的映射，以至于第一个圆的 N 个点被映射到第二个圆的一点(绕第二个圆回转 N 次)。由于这个原因，整数 N 被称作绕数，并且每个同伦类都由它的绕数刻画，其形式为 $N=-\mathrm{i}\int_0^{2\pi}(\mathrm{d}\theta/2\pi)[(1/\alpha(\theta))(\mathrm{d}\alpha/\mathrm{d}\theta)]$。因此，$\Pi_1(\mathrm{U}(1))=\Pi_1(S^1)=Z$，其中 Z 表示一个整数加法群。任何绕数的映射都能通过采用 $\alpha^{(1)}(\theta)=\exp(\mathrm{i}\theta)$ 的幂而获得。

通过让 $X=S^n$(n 维球面)可以推广这一讨论。$S^n\rightarrow S^m$ 的映射类形成一个群，叫做 S^m 的第 n 同伦群，由 $\Pi_n(S^m)$ 标记。相似于 $\Pi_1(S^1)=Z$，我们有 $\Pi_n(S^n)=Z$。也就是说，映射 $S^n\rightarrow S^n$ 也用一个 n 维球面覆盖另一个球面的次数来分类。在无限远的距离上，所有点都是同一的寻常空间 R^3 与 S^3 等价。由于 SU(2)群中的任何元素 M 能被写为 $M=a+\mathrm{i}\boldsymbol{b}\cdot\boldsymbol{\tau}$，其中 $\boldsymbol{\tau}$ 是泡利矩阵，a 和 $\boldsymbol{b}$ 满足 $a^2+\boldsymbol{b}^2=1$，因此 SU(2)群元素的流形也与 S^3 拓扑等价。因此，$\Pi_3(\mathrm{SU}(2))=\Pi_3(S^3)=Z$。由于 S^3 到一个任意群 SU(N)的任何连

续映射，能被连续地变形为一个到 SU(N)的亚群 SU(2)的映射，因此一般而言，紧致非阿贝尔规范群[包括 SU(N)]也拥有相同的结果，即 Π_3（紧致非阿贝尔规范群）= Z。

对于提供映射 $S^3 \to G$ 的规范变换来说，可以表明，其绕数由 $N = (1/24\pi^2)\int d\boldsymbol{r}\,\varepsilon^{ijk}\,\mathrm{tr}[\mathrm{U}^{-1}(\boldsymbol{r})\partial_i \mathrm{U}(\boldsymbol{r})\mathrm{U}^{-1}(\boldsymbol{r})\partial_j \mathrm{U}(\boldsymbol{r})\mathrm{U}^{-1}(\boldsymbol{r})\partial_k \mathrm{U}(\boldsymbol{r})]$ 给出，这由规范势的大距离性质确定。只当群 G 是 U(1)时，S^3 到 U(1)的每个映射才连续可变形为与 $N = 0$ 相对应的常数映射。在这种情形中，规范变换被叫做同伦平凡的。所有不可变形到单位矩阵的同伦非平凡规范变换，被叫做有限大的或大的，而平凡规范变换被叫做无限小的或小的。 [376]

参考书目

缩写词

AJP	*American Journal of Physics*
AP	*Annals of Physics*(*New York*)
CMP	*Communications in Mathematical Physics*
DAN	*Doklady Akademii Nauk SSSR*
EA	The Einstein Archives in the Hebrew University of Jerusalem
JETP	*Soviet Physics*, *Journal of Experimental and Theoretical Physics*
JMP	*Journal of Mathematical Physics*(*New York*)
NC	*Nuovo Cimento*
NP	*Nuclear Physics*
PL	*Physics Letters*
PR	*Physical Review*
PRL	*Physical Review Letters*
PRS	*Proceedings of the Royal Society of London*
PTP	*Progress of Theoretical Physics*
RMP	*Reviews of Modern Physics*
ZP	*Zeitschrift für Physik*

Abers, E., Zachariasan, F. and Zemach, C. (1963). 'Origin of internal symmetries', *PR*, **132**:1831—1836.

Abrams, G.S. *et al*. (1974). 'Discovery of a second narrow resonance', *PRL*, **33**:1453—1454.

Achinstein, P. (1968). *Concepts of Science*: *A Philosophical Analysis* (The Johns Hopkins University Press, Baltimore).

Adler, S. L. (1964). 'Tests of the conserved vector current and partially

conserved axial-vector current hypotheses in high-energy neutrino reactions', *PR*, **B135**:963—966.

Adler, S. L. (1965). 'Sum-rules for axial-vector coupling-constant renormalization in beta decay', *PR*, **B140**:736—747.

Adler, S. L. and Dashen, R. F. (1968). *Current Algebra and Applications to Particle Physics* (Benjamin, New York).

Adler, S. L. (1969). 'Axial-vector vertex in spinor electrodynamics', *PR*, **177**:2426—2438.

Adler, S. L. and Bardeen, W. (1969). 'Absence of higher-order corrections in the anomalous axial-vector divergence equation', *PR*, **182**: 1517—1532.

Adler, S. L. and Wu-Ki Tung(1969). 'Breakdown of asymptotic sum rules in perturbation theory', *PRL*, **22**:978—981.

Adler, S. L. (1970). 'π^0 decay', in *High-Energy Physics and Nuclear Structure*, ed. Devons, S. (Plenum, New York), 647—655.

Adler, S. L. (1991). Taped interview at Princeton, 5 Dec. 1991.

Aharonov, Y. and Bohm, D. (1959). 'Significance of electromagnetic potentials in quantum theory', *PR*, **115**:485—491.

Ambarzumian, V. and Iwanenko, D. (1930). 'Unobservable electrons and β-rays', *Compt. Rend. Acad. Sci. Paris*, **190**:582—584.

Anderson, P. W. (1952). 'An approximate quantum theory of the antiferromagnetic ground state', *PR*, **86**:694—701.

Anderson, P. W. (1958a). 'Coherent excited states in the theory of superconductivity: Gauge invariance and the Meissner effect', *PR*, **110**: 827—835.

Anderson, P. W. (1958b). 'Random phase approximation in the theory of super-conductivity', *PR*, **112**:1900—1916.

Anderson, P. W. (1963). 'Plasmons, gauge invariance, and mass', *PR*, **130**:439—442.

Appelquist, T. and Carazzone, J. (1975). 'Infrared singularities and massive fields', *PR*, **D11**:2856—2861.

Appelquist, T. and Politzer, H. D. (1975). 'Heavy quarks and e^+e^- annihilation', *PRL*, **34**:43—45.

Arnison, G. *et al.* (1983a). 'Experimental observation of isolated large transverse energy electrons with asociated missing energy at $\sqrt{s}$ = 540 Gev', *PL*, **122B**:103—116.

Arnison, G. *et al.* (1983b). 'Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider', *PL*, **126B**: 398—410.

Arnowitt, R., Friedman, M. H. and Nath, P. (1968). 'Hard meson analysis of photon decays of π^0, η and vector mesons', *PL*, **27B**:657—659.

Aubert, J. J. *et al.* (1974). 'Observation of a heavy particle J', *PRL*, **33**: 1404—1405.

Augustin, J.-E. *et al.* (1974). 'Discovery of a narrow resonance in e^+e^- annihilation', *PRL*, **33**:1406—1407.

Bagnaia, P. *et al*. (1983). 'Evidence for $Z^0 \to e^+ e^-$ at the CERN $p^+ p^-$ collider', *PL*, **129B**:130—140.

Baker, M. and Glashow, S. L. (1962). 'Spontaneous breakdown of elementary particle symmetries', *PR*, **128**:2462—2471.

Balzer, W., Pearce, D. A. and Schmidt, H.-J. (1984). *Reduction in Science* (Reidel, Dordrecht).

Bardeen, J. (1955). 'Theory of the Meissner effect in superconductors', *PR*, **97**:1724—1725.

Bardeen, J. (1957). 'Gauge invariance and the energy gap model of superconductivity', *NC*, **5**:1766—1768.

Bardeen, J., Cooper, L. N. and Schrieffer, J. R. (1957). 'Theory of superconductivity', *PR*, **108**:1175—1204.

Bardeen, W. (1969). 'Anomalous Ward identities in spinor field theories', *PR*, **184**:1848—1859.

Bardeen, W., Fritzsch, H. and Gell-Mann, M. (1973). 'Light cone current algebra, π^0 decay and $e^+ e^-$ annihilation', in *Scale and Conformal Symmetry in Hadron Physics*, ed. Gatto, R. (Wiley, New York), 139—153.

Bardeen, W. (1985). 'Gauge anomalies, gravitational anomalies, and superstrings', talk presented at the INS International Symposium, Tokyo, August 1985; Fermilab preprint: Conf. 85/110-T.

Barnes, B. (1977). *Interests and the Growth of Knowledge* (Routledge and Kegan Paul, London).

Belavin, A. A., Polyakov, A. M., Schwartz, A. and Tyupkin, Y. (1975). 'Pseudoparticle solutions of the Yang-Mills equations', *PL*, **59B**: 85—87.

Bell, J. S. (1967a). 'Equal-time commutator in a solvable model', *NC*, **47A**:616—625.

Bell, J. S. (1967b). 'Current algebra and gauge invariance', *NC*, **50A**: 129—134.

Bell, J. S. and Berman, S. M. (1967). 'On current algebra and CVC in pion beta-decay', *NC*, **47A**:807—810.

Bell, J. S. and Jackiw, R. (1969). 'A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model', *NC*, **60A**:47—61.

Bell, J. S. and van Royen, R. P. (1968). 'Pion mass difference and current algebra', *PL*, **25B**:354—356.

Bell, J. S. and Sutherland, D. G. (1968). 'Current algebra and $\eta \to 3\pi$', *NP*, **B4**:315—325.

Bernstein, J., (1968). *Elementary Particles and Their Currents* (Freeman, San Francisco).

Bernstein, J., Fubini, S., Gell-Mann, M. and Thirring, W. (1960). 'On the decay rate of the charged pion', *NC*, **17**:758—766.

Bethe, H. A. (1947). 'The electromagnetic shift of energy levels', *PR*, **72**:339—341.

Bjorken, J. D. and Glashow, S. L. (1964). 'Elementary particles and SU(4)', *PL*, **11**:255—257.

Bjorken, J. D. (1966). 'Applications of the chiral U(6) ⊗ U(6) algebra of current densities', *PR*, **148**:1467—1478.
Bjorken, J. D. (1969). 'Asymptotic sum rules at infinite momentum', *PR*, **179**:1547—1553.
Bjorken, J. D. and Paschos, E. A. (1969). 'Inelastic electron-proton and γ-proton scattering and the structure of the nucleon', *PR*, **185**: 1975—1982.
Bjorken, J. D. and Paschos, E. A. (1970). 'High-energy inelastic neutrino-nucleon interactions', *PR*, **D1**:3151—3160.
Blankenbecler, R., Cook, L. F. and Goldberger, M. L. (1962). 'Is the photon an elementary particle?' *PR*, **8**:463—465.
Blankenbecler, R., Coon, D. D. and Roy, S. M. (1967). 'S-matrix approach to internal symmetry', *PR*, **156**:1624—1636.
Blatt, J. M., Matsubara, T. and May, R. M. (1959). 'Gauge invariance in the theory of superconductivity', *PTP*, **21**:745—757.
Bloom, E. D. *et al*. (1969). 'High energy inelastic e-p scattering at 6° and 10°', *PRL*, **23**:930—934.
Bloor, D. (1976). *Knowledge and Social Imagery* (Routledge and Kegan Paul, London).
Bludman, S. A. (1958). 'On the universal Fermi interaction', *NC*, **9**:433—444.
Bogoliubov, N. N. (1958). 'A new method in the theory of superconductivity', *JETP*, **34**(7):41—46, 51—55.
Bohr, N. (1912). In *On the Constitution of Atoms and Molecules*, ed. Rosenfeld, L. (Benjamin, New York, 1963), p. xxxii.
Bohr, N. (1913a, b, c). 'On the constitution of atoms and molecules. I, II, III', *Philos. Mag.*, **26**:1—25, 476—502, 857—875.
Bohr, N. (1918). 'On the quantum theory of line-spectra', *Kgl. Dan. Vid. Selsk. Skr. Nat-Mat. Afd.* series 8, vol. 4, number 1, Part I-III (Høst, Copenhagen).
Bohr, N., Kramers, H. A. and Slater, J. C. (1924). 'The quantum theory of radiation', *Philos. Mag.*, **47**:785—802.
Bohr, N. (1927). 'Atomic theory and wave mechanics', *Nature*, **119**:262.
Bohr, N. (1928). 'The quantum postulate and the recent development of atomic theory' (a version of the Como lecture given in 1927), *Nature*, **121**:580—590.
Bohr, N. (1930). 'Philosophical aspects of atomic theory', *Nature*, **125**: 958.
Boltzmann, L. (1888). Quoted from R. S. Cohen's 'Dialectical materialism and Carnap's logical empiricism' in *The Philosophy of Rudolf Carnap*, ed. Schilpp, P. A. (Open Court, LaSalle, 1963), 109.
Bopp, F. (1940). 'Eine lineare theorie des elektrons', *Ann. Phys.* **38**: 345—384.
Born, M. (1924). 'Über Quantenmechanik', *ZP*, **26**:379—395.
Born, M. and Jordan, P. (1925). 'Zur Quantenmechanik', *ZP*, **34**: 858—888.

Born, M., Heisenberg, W. and Jordan, P. (1926). 'Zur Quantenmechnik. II', *ZP*, **35**: 557—615.

Born, M. (1926). 'Zur Quantenmechanik der Stossvorgange', *ZP*, **37**: 863—867.

Born, M. (1949). *Natural Philosophy of Cause and Chance* (Oxford University Press, Oxford).

Born, M. (1956). *Physics in My Generation* (Pergamon, London).

Bose, S.N. (1924). 'Plancks Gesetz und Lichtquantenhypothese', *ZP*, **26**: 178—181.

Bouchiat, C., Iliopoulos, J. and Meyer, Ph. (1972). 'An anomaly-free version of Weinberg's model', *PL*, **38B**: 519—523.

Boulware, D. (1970). 'Renormalizability of massive non-Abelian gauge fields: a functional integral approach', *AP*, **56**: 140—171.

Boyd, R. N. (1979). 'Metaphor and theory change', in *Metaphor and Thought*, ed. Ortony, A. (Cambridge University Press, Cambridge).

Boyd, R. N. (1983). 'On the current status of scientific realism', *Erkenntnis*, **19**: 45—90.

Brandt, R.A. (1967). 'Derivation of renormalized relativistic perturbation theory from finite local field equations', *AP*, **44**: 221—265.

Brandt, R.A. and Orzalesi, C.A. (1967). 'Equal-time commutator and zero-energy theorem in the Lee model', *PR*, **162**: 1747—1750.

Brans, C. and Dicke, R. H. (1961). 'Mach's principle and a relativistic theory of gravitation', *PR*, **124**: 925—935.

Brans, C. (1962). 'Mach's principle and the locally measured gravitational constant in general relativity', *PR*, **125**: 388—396.

Braunbeck, W. and Weinmann, E. (1938). 'Die Rutherford-Streuung mit Berücksichtigung der Austrahlung', *ZP*, **110**: 360—372.

Breidenbach, M. *et al*. (1969). 'Observed behavior of highly inelastic electron-proton scattering', *PRL*, **23**: 935—999.

Bridgeman, P. W. (1927). *The Logic of Modern Physics* (Macmillan, New York).

Bromberg, J. (1976). 'The concept of particle creation before and after quantum mechanics', *Historical Studies in the Physical Sciences*, **7**: 161—191.

Brown, H.R. and Harré, R. (1988). *Philosophical Foundations of Quantum Field Theory* (Clarendon, Oxford).

Brown, L.M. and Hoddeson, L. (eds.) (1983). *The Birth of Particle Physics* (Cambridge University Press, Cambridge).

Brown, L.M. and Rechenberg, H. (1988). 'Landau's work on quantum field theory and high energy physics (1930—1961)', Max Planck Institute, Preprint, MPI-PAE/Pth 42/88(July 1988).

Brown, L. M., Dresden, M. and Hoddeson, L. (eds.) (1989). *Pions to Quarks* (Cambridge University Press, Cambridge).

Brown, L. M. and Cao, T. Y. (1991). 'Spontaneous breakdown of symmetry: its rediscovery and integration into quantum field theory', *Historical Studies in the Physical and Biological Sciences* **21**: 211—235.

Bucella, F., Veneziano, G., Gatto, R. and Okubo, S. (1966). 'Necessity of additional unitary-antisymmetric q-number terms in the commutators of spatial current components', *PR*, **149**:1268—1272.

Buckingham, M. J. (1957). 'A note on the energy gap model of superconductivity', *NC*, **5**:1763—1765.

Burtt, E. A. (1932). *The Metaphysical Foundations of Modern Physical Science* (Doubleday, Garden City). (First edition in 1924, revised version first appeared in 1932.)

Cabibbo, N. (1963). 'Unitary symmetry and leptonic decays', *PRL*, **10**: 531—533.

Cabrera, B. (1982). 'First results from a superconductive detector for moving magnetic monopoles', *PRL*, **48**:1378—1381.

Callan, C. and Gross, D. (1968). 'Crucial test of a theory of currents', *PRL*, **21**:311—313.

Callan, C. and Gross, D. (1969). 'High-energy electroproduction and the constitution of the electric current', *PRL*, **22**:156—159.

Callan, C. (1970). 'Bjorken scale invariance in scalar field theory', *PR*, **D2**:1541—1547.

Callan, C., Coleman, S. and Jackiw, R. (1970). 'A new improved energy-momentum tensor', *AP*, **59**:42—73.

Callan, C. (1972). 'Bjorken scale invariance and asymptotic behavior', *PR*, **D5**:3202—3210.

Callan, C. and Gross, D. (1973). 'Bjorken scaling in quantum field theory', *PR*, **D8**:4383—4394.

Callan, C., Dashen, R. and Gross, D. (1976). 'The structure of the vacuum', *PL*, **63B**:334—340.

Carmeli, M. (1976). 'Modified gravitational Lagrangian', *PR*, **D14**:1727.

Carmeli, M. (1977). 'SL_{2C} conservation laws of general relativity', *NC*, **18**:17—20.

Cantor, G. N. and Hodge, M. J. S. (eds.) (1981). *Conceptions of Ether* (Cambridge University Press, Cambridge).

Cao, T. Y. (1991). 'The Reggeization program 1962—1982: attempts at reconciling quantum field theory with S-matrix theory', *Archive for History of Exact Sciences*, **41**:239—283.

Cao, T. Y. (1993). 'What is meant by social constructivism? — A critical exposition', a talk given at the Dibner Institute, MIT, on 19 October 1993.

Cao, T. Y. and Schweber, S. S. (1993). 'The conceptual foundations and the philosophical aspects of renormalization theory', *Synthese*, **97**:33—108.

Capps, R. H. (1963). 'Prediction of an interaction symmetry from dispersion relations, *PRL*, **10**:312—314.

Capra, F. (1979). 'Quark physics without quarks: a review of recent developments in S-matrix theory', *AJP*, **47**:11—23.

Capra, F. (1985). 'Bootstrap physics: a conversation with Geoffrey Chew', in *A Passion for Physics: Essays in Honour of Geoffrey Chew*, eds. De Tar, C., Finkelstein, J. and Tan, C. I. (Taylor and Francis,

Philadelphia), 247—286.
Carnap, R. (1929). *Der Logisches Aufbau der Welt* (Schlachtensee Weltkreis-Verlag, Berlin).
Carnap, R. (1956). *Meaning and Necessity*, enlarged edition (University of Chicago Press, Chicago).
Cartan, E. (1922). 'Sur une géneralisation de la notion de courbure de Riemann et les éspaces à torsion', *Compt. Rend. Acad. Sci. Paris*, **174**:593—595.
Case, K. M. (1949). 'Equivalence theorems for meson-nucleon coupling', *PR*, **76**:1—14.
Cassidy, D. C. and Rechenberg, H. (1985). 'Biographical data, Werner Heisenberg (1901—1976)', in W. Heisenberg, *Collected Works*, eds. Blum, W. *et al.*, Series A, part I (Berlin), 1—14.
Castillejo, L., Dalitz, R. H. and Dyson, F. J. (1956). 'Low's scattering equation for the charged and neutral scalar theories', *PR*, **101**: 453—458.
Cauchy, A. L. (1828). *Exercise de Mathématiques*, **3**:160.
Cauchy, A. L. (1830). 'Mémoire sur la théorie de la lumière', *Mém. de l'Acad.*, **10**:293—316.
Chadwick, J. (1914). 'Intensitätsvertieilung im magnetischen Spektrum der β-strahlen von Radium B + C', *Ber. Deut. Phys. Gens.* **12**:383—391.
Chadwick, J. (1932). 'Possible existence of a neutron', *Nature*, **129**:312.
Chambers, R. G. (1960). 'Shift of an electron interference pattern by enclosed magnetic flux', *PRL*. **5**:3—5.
Chandler, C. (1968). 'Causality in S-matrix theory', *PR*. **174**: 1749—1758.
Chandler, C., and Stapp, H. P. (1969). 'Macroscopic causality and properties of scattering amplitudes', *JMP*. **10**:826—859.
Chandrasekhar, S. (1931). 'The maximum mass of ideal white dwarfs', *Astrophys. J.*, **74**:81.
Chandrasekhar, S. (1934). 'Stellar configurations with degenerate cores', *Observatory*, **57**:373—377.
Chandrasekhar, S. (1935). *Observatory*, **58**:38.
Chanowitz, M. S., Furman, M. A. and Hinchliffe, Z. (1978). 'Weak interactions of ultra heavy fermions', *PR*, **B78**:285—289.
Charap, J. M. and Fubini, S. (1959). 'The field theoretic definition of the nuclear potential- I', *NC*, **14**:540—559.
Chew, G. F. (1953a). 'Pion-nucleon scattering when the coupling is weak and extended', *PR*, **89**:591—593.
Chew, G. F. (1953b). 'A new theoretical approach to the pion-nucleaon interaction', *PR*, **89**:904.
Chew, G. F. and Low, F. E. (1956). 'Effective-range approach to the low-energy *p*-wave pion-nucleon interaction', *PR*, **101**:1570—1579.
Chew, G. F., Goldberger, M. L., Low, F. E. and Nambu, Y. (1957a). 'Application of dispersion relations to low-energy meson-nucleon scattering', *PR*, **106**:1337—1344.

Chew, G. F., Goldberger, M. L., Low, F. E. and Nambu, Y. (1957b). 'Relativistic dispersion relation approach to photomeson production', *PR*, **106**: 1345—1355.

Chew, G. F. and Frautschi, S. C. (1960). 'Unified approach to high- and low-energy strong interactions on the basis of the Mandelstam representation', *PRL*, **5**: 580—583.

Chew, G. F. and Mandelstam, S. (1960). 'Theory of low energy pion-pion interaction', *PR*, **119**: 467—477.

Chew, G. F. (1961). *S-Matrix Theory of Strong Interactions* (Benjamin, New York).

Chew, G. F. and Frautschi, S. C. (1961a). 'Dynamical theory for strong interactions at low momentum transfer but arbitrary energies', *PR*, **123**: 1478—1486.

Chew, G. F. and Frautschi, S. C. (1961b). 'Potential scattering as opposed to scattering associated with independent particles in the S-matrix theory of strong interactions', *PR*, **124**: 264—268.

Chew, G. F. and Frautschi, S. C. (1961c). 'Principle of equivalence for all strongly interacting particles within the S-matrix framework', *PRL*, **7**: 394—397.

Chew, G. F. and Frautschi, S. C. (1962). 'Regge trajectories and the principle of maximum strength for strong interactions', *PRL*, **8**: 41—44.

Chew, G. F. (1962a). 'S-matrix theory of strong interactions without elementary particles', *RMP*, **34**: 394—401.

Chew, G. F. (1962b). 'Reciprocal bootstrap relationship of the nucleon and the(3, 3) resonance', *PRL*, **9**: 233—235.

Chew, G. F. (1962c). 'Strong interaction theory without elementary particles', in *Proceedings of the 1962 International Conference on High Energy Physics at CERN*, ed. Prentki, J. (CERN, Geneva), 525—530.

Chew, G. F. (1989). 'Particles as S-matrix poles: hadron democracy', in *Pions to Quarks: Particle Physics in the 1950s*, eds. Brown, L. M., Dresden, M. and Hoddeson, L. (Cambridge University Press, Cambridge), 600—607.

Chihara, C. (1973). *Ontology and the Vicious Circle Principle* (Cornell University Press, Ithaca).

Cho, Y. M. (1975). 'Higher-dimensional unifications of gravitation and gauge theories', *JMP*, **16**: 2029—2035.

Cho, Y. M. (1976). 'Einstein Lagrangian as the translational Yang-Mills Lagrangian', *PR*, **D14**: 2521—2525.

Christoffel, E. B. (1869a). 'Ueber die Transformation der homogenen Differential-ausdrücke zweiten Grades', *J. r. angew. Math.* **70**: 46—70.

Christoffel, E. B. (1869b). 'Ueber ein die Transformation hamogen Differential-ausdrücke zweiten Grades betreffendes Theorem', *J. r. angew. Math.* **70**: 241—245.

Clifford, W. K. (1876). 'On the space-theory of matter', in *Mathematical Papers*, ed. Tucker, R. (Macmillan, London).

Coleman, S. (1965). 'Trouble with relativistic SU(6)', *PR*, **B138**: 1262—1267.
Coleman, S. and Mandula, J. (1967). 'All possible symmetries of the S matrix', *PR*, **B159**:1251—1256.
Coleman, S. and Jackiw, R. (1971). 'Why dilatation generators do not generate dilatations', *AP*, **67**:552—598.
Coleman, S. and Gross, D. (1973). 'Price of asymptotic freedom', *PRL*, **31**:851—854.
Coleman, S. and Weinberg, E. (1973). 'Radiative corrections as the origin of spontaneous symmetry breaking', *PR*, **D7**:1888—1910.
Coleman, S. (1977). 'The use of instantons', a talk later published in *The Ways in Subnuclear Physics*, ed. Zichichi, A. (Plenum, New York, 1979).
Coleman, S. (1979). 'The 1979 Nobel Prize in physics', *Science*, **206**: 1290—1292.
Coleman, S. (1985). *Aspects of Symmetry* (Cambridge University Press, Cambridge).
Collins, C. B. and Hawking, S. W. (1973a). 'The rotation and distortion of the universe' *Monthly Notices of the Royal Astronomical Society*, **162**: 307—320.
Collins, C. B. and Hawking, S. W. (1973b). 'Why is the universe isotropic?', *Astrophys. J.* **180**:317—334.
Collins, J. C., Wilczek and Zee, A. (1978). 'Low-energy manifestations of heavy particles: application to the neutral current', *PR*. **D18**:242—247.
Collins, J. C. (1984). *Renormalization* (Cambridge University Press, Cambridge).
Collins, P. D. B. (1977). *An Introduction to Regge Theory and High Energy Physics* (Cambridge University Press, Cambridge).
Compton, A. (1923a). 'Total reflection of X-rays', *Philos. Mag.* **45**: 1121—1131.
Compton, A. (1923b). 'Quantum theory of the scattering of X-rays by light elements', *PR*, **21**:483—502.
Coster, J. and Stapp, H. P. (1969). 'Physical-region discontinuity equations for many-particle scattering amplitudes. I', *JMP*, **10**:371—396.
Coster, J. and Stapp, H. P. (1970a). 'Physical-region discontinuity equations for many-particle scattering amplitudes. II', *JMP*, **11**:1441—1463.
Coster, J. and Stapp, H. P. (1970b). 'Physical-region discontinuity equations', *JMP*, **11**:2743—2763.
Creutz, M. (1981). 'Roulette wheels and quark confinement', *Comments on Nuclear and Particle Physics*, **10**:163—173.
Crewther, R., Divecchia, P., Veneziano, G. and Witten, E. (1979). 'Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics', *PL*, **88B**:123—127.
Curie, P. (1894). 'Sur la symetrie dans les phenomenes physiques, symetrie d'un champ electrique et d'un champ magnetique', *J. Phys. (Paris)*, **3**: 393—415.

Cushing, J. T. (1986). 'The importance of Heisenberg's S-matrix program for the theoretical high-energy physics of the 1950's', *Centaurus*, **29**: 110—149.

Cushing, J. T. (1990). *Theory Construction and Selection in Modern Physics: The S-Matrix Theory* (Cambridge University Press, Cambridge).

Cutkosky, R. E. (1960). 'Singularities and discontinuities of Feynman amplitudes', *JMP*, **1**: 429—433.

Cutkosky, R. E. (1963a). 'A model of baryon states', *AP*, **23**: 415—438.

Cutkosky, R. E. (1963b). 'A mechanism for the induction of symmetries among the strong interactions', *PR*, **131**: 1888—1890.

Cutkosky, R. E. and Tarjanne, P. (1963). 'Self-consistent derivations from unitary symmetry', *PR*, **132**: 1354—1361.

Dancoff, S. M. (1939). 'On radiative corrections for electron scattering', *PR*, **55**: 959—963.

Daniel, M. and Viallet, C. M. (1980). 'The geometrical setting of gauge theories of the Yang-Mills type', *RMP*, **52**: 175—197.

de Broglie, L. (1923a). 'Ondes et quanta', *Compt. Rend. Acad. Sci. Paris*, **177**: 507—510.

de Broglie, L. (1923b). 'Quanta de lumiere, diffraction et interferences', *Compt. Rend. Acad. Sci. Paris*, **177**: 548—550.

de Broglie, L. (1926). 'The new undulatory mechanics', *Compt. Rend. Acad. Sci. Paris*, **183**: 272—274.

de Broglie, L. (1927a). 'Possibility of relating interference and diffraction phenomena to the theory of light quanta', *Compt. Rend. Acad. Sci. Paris*, **183**: 447—448.

de Broglie, L. (1927b). 'La mécanique ondulatoire et la structure atomique de la matiére et du rayonnement', *J. Phys. Radium* **8**: 225—241.

de Broglie, L. (1960). *Non-linear Wave Mechanics: A Causal Interpretation*, trans. Knobel, A. J. and Miller, J. C. (Elsevier, Amsterdam).

de Broglie, L. (1962). *New Perspectives in Physics*, trans. Pomerans, A. J. (Oliver and Boyd, Edinburgh).

Debye, P. (1910a). Letter to Sommerfeld, 2 March 1910.

Debye, P. (1910b). 'Der Wahrscheinlichkeitsbegriff in der Theorie der Strahlung', *Ann. Phys.*, **33**: 1427—1434.

Debye, P. (1923). 'Zerstreuung von Röntgenstrahlen und quantentheorie', *Phys. Z.*, **24**: 161—166.

Demopoulos, W. and Friedman, M. (1985). 'Critical notice: Bertrand Russell's *The Analysis of Matter*: its historical context and contemporary interest', *Philos. Sci.*, **52**: 621—639.

Deser, S. (1970). 'Self-interaction and gauge invariance', *General Relativity and Gravitation*, **1**: 9.

de Sitter, W. (1916a). 'On the relativity of rotation in Einstein's theory', *Proceedings of the Section of Science (Koninklijke Akademie van Wetenschappen te Amsterdam)*, **19**: 527—532.

de Sitter, W. (1916b). 'On Einstein's theory of gravitation and its astronomical consequences. I, II', *Monthly Notices of the Royal*

Astronomical Society, **76**:699—738; **77**:155—183.
de Sitter, W. (1917a). 'On the relativity of inertia: remarks concerning Einstein's latest hypothesis', *Proceedings of the Section of Science* (*Koninklijke Akademie van Wetenschappen te Amsterdam*), **19**: 1217—1225.
de Sitter, W. (1917b). 'On the curvature of space', *Proceedings of the Section of Science* (*Koninklijke Akademie van Wetenschappen te Amsterdam*), **20**:229—242.
de Sitter, W. (1917c). 'On Einstein's theory of gravitation and its astronomical consequences. Third paper', *Monthly Notices of the Royal Astronomical Society*, **78**:3—28.
de Sitter, W. (1917d). 'Further remarks on the solutions of the field equations of Einstein's theory of gravitation', *Proceedings of the Section of Science* (*Koninklijke Akademie van Wetenschappen te Amsterdam*), **20**: 1309—1312.
de Sitter, W. (1917e). Letter to Einstein, 1 April 1917, EA:20—551.
de Sitter, W. (1920). Letter to Einstein, 4 November 1920, EA:20—571.
de Sitter, W. (1931). 'Contributions to a British Association discussion on the evolution of the universe', *Nature*, **128**:706—709.
DeWitt, B. S. (1964). 'Theory of radiative corrections for non-Abellian gauge fields', *PRL*, **12**:742—746.
DeWitt, B. S. (1967a). 'Quantum theory of gravity. Ⅰ. The canonical theory', *PR*, **160**:1113—1148.
DeWitt, B. S. (1967b). 'Quantum theory of gravity. Ⅱ. The manifestly covariant theory', *PR*, **162**:1195—1239.
DeWitt, B. S. (1967c). 'Quantum theory of gravity. Ⅲ. Applications of the covariant theory', *PR*, **162**:1239—1256.
Dirac, P. A. M. (1925). 'The fundamental equations of quantum mechanics', *PRS*, **A109**:642—653.
Dirac, P. A. M. (1926a). 'Quantum mechanics and a preliminary investigation of the hydrogen atom', *PRS*, **A110**:561—579.
Dirac, P. A. M. (1926b). 'On the theory of quantum mechanics', *PRS*, **A112**:661—677.
Dirac, P. A. M. (1927a). 'The physical interpretation of the quantum dynamics', *PRS*, **A113**:621—641.
Dirac, P. A. M. (1927b). 'The quantum theory of emission and absorption of radiation', *PRS*, **A114**:243—265.
Dirac, P. A. M. (1927c). 'The quantum theory of dispersion', *PRS*, **A114**: 710—728.
Dirac, P. A. M. (1928a). 'The quantum theory of the electron', *PRS*, **A117**:610—624.
Dirac, P. A. M. (1928b). 'The quantum theory of the electron. Part Ⅱ', *PRS*, **A118**:351—361.
Dirac, P. A. M. (1930a). 'A theory of electrons and protons', *PRS*, **A126**: 360—365.
Dirac. P. A. M. (1930b). *The Principles of Quantum Mechanics* (Clarendon,

Oxford).

Dirac, P. A. M. (1931). 'Quantized singularities in the electromagnetic field', *PRS*, **A133**:60—72.

Dirac, P. A. M. (1932). 'Relativistic quantum mechanics', *PRS*, **A136**: 453—464.

Dirac, P. A. M. (1933). 'Théorie du positron', in *Rapport du Septième Conseil de Solvay Physique, Structure et Propriétés des noyaux atomiques (22—29 Oct. 1933)* (Gauthier-Villars, Paris), 203—212.

Dirac, P. A. M. (1934). 'Discussion of the infinite distribution of electrons in the theory of the positron', *Proc. Cambridge Philos. Soc.*, **30**:150—163.

Dirac, P. A. M. (1938). 'Classical theory of radiating electrons', *PRS*, **A167**:148—169.

Dirac, P. A. M. (1939). 'La théorie de l'électron et du champ électromagnetique', *Ann. Inst. Henri Poincaré*, **9**:13—49.

Dirac, P. A. M. (1942). 'The physical interpretation of quantum mechanics', *PRS*, **A180**:1—40.

Dirac, P. A. M. (1948). 'Quantum theory of localizable dynamic systems', *PR*, **73**:1092—1103.

Dirac, P. A. M. (1951). 'Is there an aether?' *Nature*, **168**:906—907.

Dirac, P. A. M. (1952). 'Is there an aether?' *Nature*, **169**:146 and 702.

Dirac, P. A. M. (1963). 'The evolution of the physicist's picture of nature', *Scientific American*, **208**(5):45—53.

Dirac, P. A. M. (1968). 'Methods in theoretical physics', in *Special Supplement of IAEA Bulletin* (IAEA, Vienna, 1969), 21—28.

Dirac, P. A. M. (1969a). 'Can equations of motion be used?', in *Coral Gables Conference on Fundamental Interactions at High Energy, Coral Gables, 22—24 Jan. 1969* (Gordon and Breach, New York), 1—18.

Dirac, P. A. M. (1969b). 'Hopes and fears', *Eureka*, **32**:2—4.

Dirac, P. A. M. (1973a). 'Relativity and quantum mechanics', in *The Past Decades in Particle Theory*, eds. Sudarshan, C. G. and Neéman, Y. (Gordon and Breach, New York), 741—772.

Dirac, P. A. M. (1973b). 'Development of the physicist's conception of nature', in *The Physicist's Conception of Nature*, ed. Mehra, J. (Reidel, Dordrecht), 1—14.

Dirac, P. A. M. (1977). 'Recollections of an exciting era', in *History of Twentieth Century Physics*, ed. Weiner, C. (Academic Press, New York), 109—146.

Dirac, P. A. M. (1978). *Directions in Physics* (Wiley, New York).

Dirac, P. A. M. (1981). 'Does renormalization make sense?', in *Perturbative Quantum Chromodynamics*, eds. Duke, D. W. and Owen, J. F. (AIP Conference Proceedings No. 74, American Institute of Physics, New York), 129—130.

Dirac, P. A. M. (1983). 'The origin of quantum field theory', in *The Birth of Particle Physics*, eds. Brown, L. M. and Hoddeson, L. (Cambridge University Press, Cambridge), 39—55.

Dirac, P. A. M. (1984a). 'The future of atomic physics', *Int. J. Theor. Phys.*, **23**(8):677—681.
Dirac, P. A. M. (1984b). 'The requirements of fundamental physical theory', *Eur. J. Phys.*, **5**:65—67.
Dirac, P. A. M. (1987). 'The inadequacies of quantum field theory', in *Reminiscences about a Great Physicist: Paul Adrien Maurice Dirac*, eds. Kursunoglu, B. N. and Wigner, E. P. (Cambridge University Press, Cambridge).
Dolen, R., Horn, O. and Schmid, C. (1967). 'Prediction of Regge parameters of ρ poles from low-energy πN data', *PRL*. **19**:402—407.
Dolen, R., Horn, O. and Schmid, C. (1968). 'Finite-energy sum rules and their applications to πN charge exchange', *PR*, **166**:1768—1781.
Doran, B. G. (1975). 'Origins and consolidation of field theory in nineteenth century Britain', *Hist. Stud. Phys. Sci.*, **6**:133—260.
Dorfman, J. (1930). 'Zur Frage über die magnetischen Momente der Atomkerne', *ZP*, **62**:90—94.
Duane, W. (1923). 'The transfer in quanta of radiation momentum to matter', *Proc. Nat. Acad. Sci.*, **9**:158—164.
Duhem, P. (1906). *The Aim and Structure of Physical Theory*, (Princeton University Press, Princeton, 1954).
Dürr, H. P., Heisenberg, W., Mitter, H., Schlieder, S. and Yamazaki, K. (1959). 'Zur theorie der elementarteilchen', *Zeitschrift für Naturforschung*, **14A**:441—485.
Dürr, H. P. and Heisenberg, W. (1961). 'Zur theorie der "seltsamen" teilchen', *Zeitschrift für Naturforschung*, **16A**:726—747.
Dyson, F. J. (1949a). 'The radiation theories of Tomonaga, Schwinger and Feynman', *PR*, **75**:486—502.
Dyson, F. J. (1949b). 'The *S*-matrix in quantum electrodynamics', *PR*, **75**:1736—1755.
Dyson, F. J. (1951). 'The renormalization method in quantum electrodynamics', *PRS*, **A207**:395—401.
Dyson, F. J. (1952). 'Divergence of perturbation theory in quantum electrodynamics', *PR*, **85**:631—632.
Dyson, F. J. (1965). 'Old and new fashions in field theory', *Physics Today*, **18**(6): 21—24.
Earman, J., Glymore, C. and Stachel, J. (eds.) (1977). *Foundations of Space-Time Theories* (University of Minnesota Press, Minneapolis).
Earman, J. (1979). 'Was Leibniz a relationist?' In *Studies in Metaphysics*, eds. French, P. and Wettstein, H. (University of Minnesota Press, Minneapolis).
Earman, J. and Norton, J. (1987). 'What price space-time substantivalism? The hole story', *British Journal for the Philosophy of Science*, **38**: 515—525.
Earman, J. (1989). *World-Enough and Space-Time* (MIT Press, Cambridge, MA).
Eddington, A. S. (1916). Letter to de Sitter, 13 October 1916, Leiden

Observatory, quoted by Kerszberg(1989).
Eddington, A. S. (1918). *Report on the Relativity theory of Gravitation* (Fleetway, London).
Eddington, A. S. (1921). 'A generalization of Weyl's theory of the electromagnetic and gravitational fields', *PRS*, **A99**:104—122.
Eddington, A. S. (1923). *The Mathematical Theory of Gravitation* (Cambridge University Press, Cambridge).
Eddington, A. S. (1926). *The Internal Constitution of the Stars* (Cambridge University Press, Cambridge).
Eddington, A. S. (1930). 'On the instability of Einstein's spherical world', *Monthly Notices of the Royal Astronomical Society*, **90**:668—678.
Eddington, A. S. (1935). 'Relativistic degeneracy', *Observatory*, **58**: 37—39.
Ehrenfest, P. (1906). 'Zur Planckschen Strahlungstheorie', *Phys. Z.*, **7**: 528—532.
Ehrenfest, P. (1911). 'Welche Zuge der lichtquantenhypothese spielen in der Theorie die Warmestrahlung eine wesentliche Rolle?' *Ann. Phys.*, **36**:91—118.
Ehrenfest, P. and Kamerling-Onnes, H. (1915). 'Simplified deduction of the formula from the theory of combinations which Planck uses as the basis for radiation theory', *Proc. Amsterdam Acad.*, **23**:789—792.
Ehrenfest, P. (1916). 'Adiabatische invarianten und quantentheorie', *Ann. Phys.*, **51**:327—352.
Eichten, E., Gottfried, K., Kinoshita, T., Koght, J., Lane, K. D. and Yan, T.-M. (1975). 'Spectrum of charmed quark-antiquark bound states', *PRL*, **34**:369—372.
Einstein, A. (1905a). 'Über einen die Erzeugung und Verwandlung des lichtes betreffenden heuristischen Gesichtspunkt', *Ann. Phys.*, **17**: 132—148.
Einstein, A. (1905b). 'Die von der molekulärkinetischen Theorie der Wärme gefor-derte bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen', *Ann. Phys.*, **17**:549—560.
Einstein, A. (1905c). 'Zur Elektrodynamik bewegter Körper', *Ann. Phys.*, **17**:891—921.
Einstein, A. (1905d). 'Ist die Trägheit eines Körpers von seinem Energiegehalt abhängig?', *Ann. Phys.*, **18**:639—641.
Einstein, A. (1906a). 'Zur Theorie der Lichterzeugung und Lichtabsorption', *Ann. Phys.*, **20**:199—206.
Einstein, A. (1906b). 'Das Prinzip von der Erhaltung der Schwerpunktsbewegung und die Trägheit der Energie', *Ann. Phys.*, **20**: 627—633.
Einstein, A. (1907a). 'Die vom Relativitätsprinzip geforderte Trägheit der Energie', *Ann. Phys.*, **23**:371—384.
Einstein, A. (1907b). 'Über das Relativitätsprinzip und aus demselben gezogenen Folgerungen', *Jahrb. Radioakt. Elektron.*, **4**:411—462.
Einstein, A. (1909a). 'Zum gegenwärtigen Stand des Strhlungsproblems',

Phys. *Z*., **10**:185—193.
Einstein, A. (1909b). 'Über die Entwicklung unserer Anschauungen über das wesen und die Konstitution der Strahlung', *Phys*. *Z*., **10**:817—825.
Einstein, A. (1909c). Letter to Lorentz, 23 May 1909.
Einstein, A. (1911). 'Über den Einfluss der Schwerkraft aus die Ausbreitung des Lichtes'; *Ann*. *Phys*. **35**:898—908.
Einstein, A. (1912a). 'Lichtgeschwindigkeit und statik des Gravitationsfeldes', *Ann*. *Phys*. **38**:355—369.
Einstein, A. (1912b). 'Zur Theorie des statischen Gravitationsfeldes', *Ann*. *Phys*. **38**:443—458.
Einstein, A. and Grossmann, M. (1913). 'Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation', *Z*. *Math*. *Phys*., **62**:225—261.
Einstein, A. (1913a). 'Zum gegenwärtigen Stande des Gravitationsproblems', *Phys*. *Z*., **14**:1249—1266.
Einstein, A. (1913b). 'Physikalische Grundlagen einer gravitationstheorie', *Vierteljahrsschr*. *Naturforsch*. *Ges*. *Zürich*, **58**:284—290.
Einstein, A. (1913c). A letter to Mach, E. 25 June 1913; quoted by Holton, G. (1973).
Einstein, A. (1914a). 'Prinzipielles zur verallgeneinerten Relativitästheorie und Gravitationstheorie', *Phys*. *Z*., **15**:176—180.
Einstein, A. (1914b). 'Die formale Grundlage der allgemeinen Relativitätstheorie', *Preussische Akad*. *Wiss*. *Sitzungsber*.: 1030—1085.
Einstein, A. (1915a). 'Zur allgemeinen Relativitätstheorie', *Preussische Akad*. *Wiss*. *Sitzungsber*.: 778—786.
Einstein, A. (1915b). 'Zur allgemeinen Relativitätstheorie (Nachtrag)', *Preussische Akad*. *Wiss*. *Sitzungsber*.:799—801.
Einstein, A. (1915c). 'Erklärung der perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie', *Preussische Akad*. *Wiss*. *Sitzungsber*.: 831—839.
Einstein, A. (1915d). 'Die Feldgleichungen der Gravitation', *Preussische Akad*. *Wiss*. *Sitzungsber*.:844—847.
Einstein, A. (1915e). Letter to Ehrenfest, P., 26 Dec. 1915. EA:9—363.
Einstein, A. (1916a). Letter to Besso, M., 3 Jan. 1916, in Speziali, P. (1972).
Einstein, A. (1916b). Letter to Schwarzschild, K., 9 Jan. 1916. EA: 21—516.
Einstein, A. (1916c). 'Die Grundlage der allgemeinen Relativitätstheorie', *Ann*. *Phys*. **49**:769—822.
Einstein, A. (1916d). 'Ernst Mach', *Phys*. *Z*., **17**:101—104.
Einstein, A. (1917a). 'Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie', *Preussische Akad*. *Wiss*. *Sitzungsber*., 142—152.
Einstein, A. (1917b). 'Zur quantentheorie der strahlung', *Phys*. *Z*., **18**: 121—128.
Einstein, A. (1917c). Letter to de Sitter, 12 March 1917. EA:20—542.
Einstein, A. (1917d). Letter to de Sitter, 14 June 1917. EA:20—556.

Einstein, A. (1917e). Letter to de Sitter, 8 August 1917. EA:20—562.
Einstein, A. (1918a). 'Prinzipielles zur allgemeinen Relativitätstheorie', *Ann. Phys*. **55**:241—244.
Einstein, A. (1918b). 'Kritischen zu einer von Herrn de Sitter gegebenen Lösung der Gravitationsgleichungen', *Preussische Akad. Wiss. Sitzungsber.*, 270—272.
Einstein, A. (1918c). 'Dialog über Einwande gegen die Relativitätstheorie', *Naturwissenschafien*. **6**:197—702.
Einstein, A. (1918d). ' "Nachtrag" zu H. Weyl: "Gravitation und elektrizität" ', *Preussische Akad. Wiss. Sitzungsber.*, 478—480.
Einstein, A. (1919). 'Spielen Gravitationsfelder im Aufbau der materiellen Elementarteilchen eine wesentliche Rolle?', *Preussische Akad. Wiss. Sitzungsber.*, 349—356.
Einstein, A. (1920a). Äther und Relativitätstheorie(Springer, Berlin).
Einstein, A. (1920b). *Relativity*(Methuen).
Einstein, A. (1921a). 'Geometrie und Erfahrung', *Preussische Akad. Wiss. Sitzungsber.*, 123—130.
Einstein, A. (1921b). 'A brief outline of the development of the theory of relativity', *Nature*, 106:782—784.
Einstein, A. (1922). *The Meaning of Relativity*(Princeton University Press, Princeton).
Einstein, A. (1923a). 'Zur affinen Feldtheorie', *Preussische Akad. Wiss. Sitzungsber.*, 137—140.
Einstein, A. (1923b). 'Bietet die Feldtheorie Möglichkeiten für die Lösung des Quanten problems?', *Preussische Akad. Wiss. Sitzungsber.*, 359—364.
Einstein, A. (1923c). 'Notiz zu der Arbeit von A. Friedmann "Über die Krümmung des Raumes" ', *ZP*, **16**:228.
Einstein, A. (1924). 'Quantentheorie des einatomigen idealen Gases', *Preussische Akad. Wiss. Sitzungsber.*, 261—267.
Einstein, A. (1925a). 'Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung', *Preussische Akad. Wiss. Sitzungsber.*, 3—14.
Einstein, A. (1925b). 'Non-Euclidean geometry and physics', *Neue Rundschau*, **1**:16—20.
Einstein, A. (1927). 'The meaning of Newton and influence upon the development of theoretical physics', *Naturwissenschaften*, **15**:273—276.
Einstein, A. (1929a). 'Field: old and new', *New York Times*, 3 Feb. 1929.
Einstein, A. (1929b). 'Zur einheitlichen Feldtheorie', *Preussische Akad. Wiss. Sitzungsber.*, 2—7.
Einstein, A. (1929c). 'Professor Einstein spricht über das physikalische Raum-und Äther-Problem', *Deutsche Bergwerks-Zeitung*, 15 Dec. 1929, p.11.
Einstein, A. (1930a). 'Raum-, Feld-und Äther-Problem in der physik', in *Gesamtbericht, Zweite Weltkraftkonferenz, Berlin, 1930*, eds. Neden, F. and Kromer, C. (VDI-Verlag, Berlin), 1—5.
Einstein, A. (1930b). Letter to Weiner, A., 18 Sept. 1930; quoted by

Holton(1973).
Einstein, A. (1931). 'Zum kosmologischen problem der allgemeinen Relativitäts-theorie', *Preussische Akad. Wiss. Sitzungsber.* 235—237.
Einstein, A. (1933a). *On the Method of Theoretical Physics* (Clarendon, Oxford).
Einstein, A. (1933b). *Origins of the General Theory of Relativity* (Jackson, Glasgow).
Einstein, A. (1936). 'Physics and reality', *J. Franklin Inst.* **221**: 313—347.
Einstein, A., Infeld, L. and Hoffmann, B. (1938). 'Gravitational equations and the problem of motion', *Ann. Math.*, **39**: 65—100.
Einstein, A. (1945). 'A generalization of relativistic theory of gravitation', *Ann. Math.*, **46**: 578—584.
Einstein, A. (1948a). A letter to Barnett, L. 19 June 1948, quoted by Stachel. J. in his 'Notes on the Andover conference', in Earman *et al.* (1977), ix.
Einstein, A. (1948b). 'Generalized theory of gravitation', *RMP*, **20**: 35—39.
Einstein, A. (1949). 'Autobiographical notes', in *Albert Einstein: Philosopher-Scientist*, ed. Schilpp, P. A. (The Library of Living Philosophers, Evanston), 1—95.
Einstein, A. (1950a). 'On the generalized theory of gravitation', *Scientific American*, **182**(4): 13—17.
Einstein, A. (1950b). Letter to Viscount Samuel, 13 Oct. 1950, in *In Search of Reality*, by Samuel, V. (Blackwell, 1957), p. 17.
Einstein, A. (1952a). 'Relativity and the problem of space', appendix 5 in the 15th edition of *Relativity: The Special and the General Theory* (Methuen, London, 1954), 135—157.
Einstein, A. (1952b). Preface to the fifteenth edition of *Relativity: The Special and the General Theory* (Methuen, 1954).
Einstein, A. (1952c). Letter to Seelig, C., 8 April 1952, quoted by Holton, G. (1973).
Einstein, A. (1952d). Letter to Max Born, 12 May 1952, in *The Born-Einstein Letters* (Walker, New York, 1971)
Einstein, A. (1953). Foreword to *Concepts of Space*, by Jammer, M. (Harvard University Press, Cambridge, MA).
Einstein, A. (1954a). Letter to Pirani, F., 2 Feb. 1954.
Einstein, A. (1954b). 'Relativity theory of the non-symmetrical field', appendix 2 in *The Meaning of Relativity* (Princeton University Press, Princeton, 1955), 133—166.
Ellis, G. (1989). 'The expanding universe: a history of cosmology from 1917 to 1960', in *Einstein and the History of General Relativity*, eds. Howard, D. and Stachel, J. (Birkhäuser, Boston/Basel/Berlin), 367—431.
Elsasser, W. (1925). 'Bemerkunggen zur Quantenmechanik Elektronen', *Naturwissenschaften*, **13**: 711.

Englert, F. and Brout, R. (1964). 'Broken symmetry and the mass of gauge vector mesons', *PRL*, **13**:321—323.
Englert, F. and Brout, R. and Thiry, M. F. (1966). 'Vector mesons in presence of broken symmetry', *NC*, **43**:244—257.
Essam, J. W. and Fisher, M. E. (1963). 'Padé approximant studies of the lattice gas and Ising ferromagnet below the critical point', *J. Chem. Phys.*, **38**:802—812.
Euler, H. (1936). 'Über die streuung von Licht an Licht nach Diracschen Theorie', *Ann. Phys.*, **21**:398—448.
Faddeev, L. D. and Popov, V. N. (1967). 'Feynman diagrams for the Yang-Mills field', *PL*, **25B**:29—30.
Faraday, M. (1844). 'A speculation touching electric conduction and the nature of matter', *Philos. Mag.*, **24**:136—144.
Feinberg, G. and Gürsey, F. (1959). 'Space-time properties and internal symmetries of strong interactions', *PR*, **114**:1153—1170.
Feldman, D. (1949). 'On realistic field theories and the polarization of the vacuum', *PR*, **76**:1369—1375.
Feldman, D. (1967). (ed.) *Proceedings of the Fifth Annual Eastern Theoretical Physics Conference* (Benjamin, New York).
Fermi, E. (1922). 'Sopra i fenomeni che avvengono in vicinanza di una linea oraria', *Accad. Lincei*, **311**:184—187, 306—309.
Fermi, E. (1929). 'Sopra l'electrodinamica quantistica. I.', *Rend. Lincei*, **9**:881—887.
Fermi, E. (1930). 'Sopra l'electrodinamica quantistica. II.', *Rend. Lincei*, **12**:431—435.
Fermi, E. (1932). 'Quantum theory of radiation', *RMP*, **4**:87—132.
Fermi, E. (1933). 'Tentativo di una teoria del l'emissione dei raggi β', *Ric. Sci.*, **4**(2): 491—495.
Fermi, E. (1934). 'Versuch einer Theorie der β-Strahlen. I', *ZP*, **88**: 161—171.
Feyerabend, P. (1962). 'Explanation, reduction and empiricism', in *Scientific Explanation, Space and Time*, eds. Feigl, H. and Maxwell, G. (University of Minnesota Press, Minneapolis), 28—97.
Feynman, R. P. (1948a). 'A relativistic cut-off for classical electrodynamics', *PR*, **74**:939—946.
Feynman, R. P. (1948b). 'Relativistic cut-off for quantum electrodynamics', *PR*, **74**:1430—1438.
Feynman, R. P. (1948c). 'Space-time approach to non-relativistic quantum mechanics', *RMP*, **20**:367—387.
Feynman, R. P. (1949a). 'The theory of positrons', *PR*, **76**:749—768.
Feynman, R. P. (1949b). 'The space-time approach to quantum electrodynamics', *PR*, **76**:769—789.
Feynman, R. P. and Gell-Mann, M. (1958). 'Theory of the Fermi interaction', *PR*, **109**:193—198.
Feynman, R. P. (1963). 'Quantum theory of gravity', *Acta Phys. Polonica*, **24**:697—722.

Feynman, R. P. and Hibbs, A. R. (1965). *Quantum Mechanics and Path Integrals*(McGraw-Hill, New York).
Feynman, R. P. (1969). 'Very high-energy collisions of hadrons', *PRL*, **23**:1415—1417.
Feynman, R. P. (1973). 'Partons', in *The Past Decade in Particle Theory*, eds. Sudarshan, C. G. and Ne'eman, Y. (Gordon and Breach, New York), 775.
Finkelstein, D. (1958). 'Past-future asymmetry of the gravitational field of a point particle', *PR*, **110**:965.
Fisher, M. E. (1964). 'Correlation functions and the critical region of simple fluids', *JMP*, **5**:944—962.
Fitzgerald, G. E. (1885). 'On a model illustrating some properties of the ether', in *The Scientific Writings*(Dublin).
Fock, V. (1926). 'Über die invariante from der Wellen-und der Bewegungsgleichungen für einen geladenen Massenpunkt', *ZP*, **39**:226—233.
Forman, P. (1971). 'Weimar culture, causality, and quantum theory, 1918—27: adaptation by German physicists and mathematicians to a hostile intellectual environment', *Historical Studies in the Physical Sciences*, **3**:1—115.
Fradkin, E. S. (1956). 'Concerning some general relations of quantum electrodynamics', *JETP*, **2**:361—363.
Fradkin, E. S. and Tyutin, I. V. (1970). 'S matrix for Yang-Mills and gravitational fields', *PR*, **D2**:2841—2857.
Frenkel, J. (1925). 'Zur elektrodynamik punktfoermiger elektronen', *ZP*, **32**:518—534.
Friedlander, F. G. (1976). *The Wave Equation on a Curved Space-time* (Cambridge University Press, Cambridge).
Friedman, A. (1922). 'Über die Krümmung des Raumes', *ZP*, **10**: 377—386.
Friedman, A. (1924). 'Über die Möglichkeit einer Welt mit konstant negativer Krümmung des Raumes', *ZP*, **21**:326—332.
Friedman, M. (1983). *Foundations of Space-Time Theories* (Princeton University Press, Princeton).
Frishman Y. (1971). 'Operator products at almost light like distances', *AP*, **66**:373—389.
Fritzsch, H., Gell-Mann, M. and Leutwyler, H. (1973). 'Advantages of the color octet gluon picture', *PL*, **47B**:365—368.
Fritzsch, H. and Minkowski, P. (1975). 'Unified interactions of leptons and hadrons', *AP*, **93**:193—266.
Fubini, S and Furlan, G. (1965). 'Renormalization effects for partially conserved currents', *Physics*, **1**:229—247.
Fukuda, H. and Miyamoto, Y. (1949). 'On the γ-decay of neutral meson', *PTP*, **4**:347—357.
Furry, W. H. and Oppenheimer, J. R. (1934). 'On the theory of the electron and positron', *PR*, **45**:245—262.
Furry, W. H. and Neuman, M. (1949). 'Interaction of meson with

electromagnetic field', *PR*, **76**:432.
Gasiorowicz, S. G., Yennie, P. R. and Suura, H. (1959). 'Magnitude of renormalization constants', *PRL*, **2**:513—516.
Gauss, C. F. (1827). 'Disquisitiones generales circa superficies curvas', *Comm. Soc. reg. Sci. Gött. cl. math.* **6**:99—146.
Gauss, C. F. (1845). Letter to Weber, 19 March 1845, quoted by Whittaker (1951), 240.
Gell-Mann, M. and Low, F. E. (1954). 'Quantum electrodynamics at small distances', *PR*, **95**:1300—1312.
Gell-Mann, M., Goldberger, M. L. and Thirring, W. (1954). 'Use of causality conditions in quantum theory', *PR*, **95**:1612—1627.
Gell-Mann, M. and Goldberger, M. L. (1954). 'The scattering of low energy photons by particles of spin 1/2', *PR*, **96**:1433—1438.
Gell-Mann, M. (1956). 'Dispersion relations in pion-pion and photon-nucleon scattering', in *High Energy Nuclear Physics, Proceedings of the Sixth Annual Rochester Conference* (Interscience, New York), sect. Ⅲ: 30—36.
Gell-Mann, M. (1958). Remarks after Heisenberg's paper, in *1958 Annual International Conference on High Energy Physics at CERN* (CERN, Geneva), 126.
Gell-Mann, M. and Levy, M. (1960). 'The axial vector current in beta decay', *NC*, **14**:705—725.
Gell-Mann, M. (1960). Remarks in *Proceedings of the 1960 Annual International Conference on High Energy Physics at Rochester* (Interscience, New York), 508.
Gell-Mann, M. (1962a). 'Symmetries of baryons and mesons', *PR*, **125**: 1067—1084.
Gell-Mann, M. (1962b). 'Applications of Regge poles', in *1962 International Conference on High Energy Physics at CERN* (CERN, Geneva), 533—542.
Gell-Mann, M. (1962c). 'Factorization of coupling to Regge poles', *PRL*, **81**:263—264.
Gell-Mann, M. and Ne'eman, Y. (1964). *The Eightfold Way* (Benjamin, New York).
Gell-Mann, M. (1964a). 'A schematic model of baryons and mesons', *PL*, **8**:214—215.
Gell-Mann, M. (1964b). 'The symmetry group of vector and axial vector currents', *Physics*, **1**:63—75.
Gell-Mann, M., Ramond, P. and Slansky, R. (1978). 'Color embeddings, charge assignments, and proton stability in unified gauge theories', *RMP*, **50**:721—744.
Gell-Mann. M. (1987). 'Particle theory from S-matrix to quarks', in *Symmetries in Physics (1600—1980)*, eds. Doncel, M. G., Hermann, A., Michael, L. and Pais, A. (Bellaterra, Barcelona), 474—497.
Gell-Mann, M. (1989). 'Progress in elementary particle theory, 1950—1964', in *Pions to Quarks*, eds. Brown, L. M., Dresden, M. and

Hoddeson, L. (Cambridge University Press, Cambridge), 694—709.
Georgi, H. and Glashow, S. L. (1974). 'Unity of all elementary particles', *PRL*, **32**:438—441.
Georgi, H., Quinn, H. R. and Weinberg, S. (1974). 'Hierarchy of interactions in unified gauge theories', *PRL*, **33**:451—454.
Georgi, H. (1989a). 'Grand unified field theories', in *The New Physics*, ed. Davies, P. (Cambridge University Press, Cambridge), 425—445.
Georgi, H. (1989b). 'Effective quantum field theories', in *The New Physics*, ed. Davies, P. (Cambridge University Press, Cambridge), 446—457.
Gershstein, S. and Zel'dovich, J. (1955). 'On corrections from mesons to the theory of β-decay', *Zh. Eksp. Teor. Fiz.* **29**: 698—699 [English trans.: *JETP*, **2**:576].
Gerstein I. and Jackiw, R. (1969). 'Anomalies in Ward identities for three-point functions', *PR*, **181**:1955—1963.
Gervais, J. L. and Lee, B. W. (1969). 'Renormalization of the σ-model(II) Fermion fields and regularization', *NP*, **B12**:627—646.
Giedymin, J. (1982). *Science and Convention* (Pergamon, Oxford).
Ginzburg, V. L. and Landau, L. D. (1950). 'On the theory of superconductivity', *Zh. Eksp. Teor. Fiz.*, **20**:1064.
Glashow, S. L. (1959). 'The renormalizability of vector meson interactions', *NP*, **10**:107—117.
Glashow, S. L. (1961). 'Partial symmetries of weak interactions', *NP*, **22**: 579—588.
Glashow, S. L., Iliopoulos, J. and Maiani, L. (1970). 'Weak interactions with leptonhadron symmetry', *PR*, **D2**:1285—1292.
Glashow, S. L. (1980). 'Toward a unified theory: threads in a tapestry', *RMP*, **52**:539—543.
Glauser, R. J. (1953). 'On the gauge invariance of the neutral vector meson theory', *PTP*, **9**:295—298.
Gödel, K. (1949). 'An example of a new type of cosmological solution of Einstein's field equations of gravitation', *RMP*, **21**:447—450.
Gödel, K. (1952). 'Rotating universes in general relativity theory', *Proc. Int. Congr. of Mathematicians, Cambridge, Mass. 1950* (American Mathematical Society, Providence), vol. I :175—181.
Goldberger, M. L. (1955a). 'Use of causality conditions in quantum theory', *PR*, **97**:508—510.
Goldberger, M. L. (1955b). 'Causality conditions and dispersion relations I. Boson fields', *PR*, **99**:979—985.
Goldberger, M. L. and Treiman, S. B. (1958a). 'Decay of the pi meson', *PR*, **110**:1178—1184.
Goldberger, M. L. and Treiman, S. B. (1958b). 'Form factors in β decay and ν capture', *PR*, **111**:354—361.
Goldhaber, G. *et al.* (1976). 'Observation in $e^+ e^-$ annihilation of a narrow state at 1865 Mev/c^2 decaying to $K\pi$ and $K\pi\pi\pi$', *PRL*, **37**:255—259.

Goldstone, J. (1961). 'Field theories with "superconductor" solutions', *NC*, **19**, 154—164.

Goldstone, J., Salam, A. and Weinberg, S. (1962) 'Broken symmetries', *PR*, **127**, 965—970.

Goto, T. and Imamura, T. (1955). 'Note on the non-perturbation-approach to quantum field theory', *PTP*, **14**:396—397.

Green, M. B. (1985). 'Unification of forces and particles in superstring theories', *Nature*, **314**:409—414.

Green, M. B. and Schwarz, J. H. (1985). 'Infinity cancellations in SO(32) superstring theory', *PL*, **151B**:21—25.

Green, M. B., Schwarz, J. H. and Witten, E. (1987). *Superstring Theory* (Cambridge University Press, Cambridge).

Gross, D. and Llewellyn Smith, C. H. (1969). 'High-energy neutrino-nucleon scattering, current algebra and partons', *NP*, **B14**:337—347.

Gross, D. and Treiman, S. (1971). 'Light cone structure of current commutators in the gluon-quark model', *PR*, **D4**:1059—1072.

Gross, D. and Jackiw, R. (1972). 'Effect of anomalies on quasi-renormalizable theories', *PR*, **D6**:477—493.

Gross, D. and Wilczek, F. (1973a). 'Ultraviolet behavior of non-Abelian gauge theories', *PRL*, **30**:1343—1346.

Gross, D. and Wilczek, F. (1973b). 'Asymptotic free gauge theories: I', *PR*, **D8**:3633—3652.

Gross, D. (1985). 'Beyond quantum field theory', in *Recent Developments in Quantum Field Theory*, eds. Ambjorn, J. Durhuus, B. J. and Petersen, J. L. (Elsevier, New York), 151—168.

Gross, D. (1992). 'Asymptotic freedom and the emergence of QCD' (a talk given at the Third International Symposium on the History of Particle Physics, June 26, 1992; Princeton Preprint PUPT 1329).

Gürsey, F. (1960a). 'On the symmetries of strong and weak interactions', *NC*, **16**:230—240.

Gürsey, F. (1960b). 'On the structure and parity of weak interaction currents', in *Proceedings of 10th International High Energy Physics Conference*, 570—577.

Gürsey, F. and Radicati, L. A. (1964). 'Spin and unitary spin independence', *PRL*, **13**:173—175.

Gürsey, F. and Sikivie, P. (1976). 'E_7 as a unitary group', *PRL*, **36**:775—778.

Guralnik, G. S., Hagen, C. R. and Kibble, T. W. B. (1964). 'Global conservation laws and massless particles', *PRL*, **13**:585—587.

Haag, R. (1958). 'Quantum field theories with composite particles and asymptotic conditions', *PR*, **112**:669—673.

Haas, A. (1910a). 'Über die elektrodynamische Bedeutung des Planck'schen strah-lungsgesetzes und über eine neue Bestimmung des elektrischen Elementarquantums und der Dimensionen des Wasserstoffatoms', *Wiener Ber*. *II*, **119**:119—144.

Haas, A. (1910b). 'Über eine neue theoretische Methode zur Bestmmung

des elektrischen Elementarquantums und des Halbmessers des Wasserstoffatoms', *Phys. Z.*, **11**:537—538.

Haas, A. (1910c). 'Der zusammenhang des Planckschen elementaren Wirkungs-quantums mit dem Grundgrössen der Elektronentheorie', *J. Radioakt.*, **7**:261—268.

Hacking, I. (1983). *Representing and Intervening* (Cambridge University Press, Cambridge).

Hamprecht, B. (1967). 'Schwinger terms in perturbation theory', *NC*, **50A**:449—457.

Han, M. Y. and Nambu, Y. (1965). 'Three-triplet model with double SU(3) symmetry', *PR*, **B139**:1006—1010.

Hanson, N. R. (1958). *Patterns of Discovery* (Cambridge University Press, Cambridge).

Hara, Y. (1964). 'Unitary triplets and the eightfold way', *PR*, **B134**: 701—704.

Harman, P. M. (1982a). *Energy, Force, and Matter. The Conceptual Development of Nineteenth-Century Physics* (Cambridge University Press, Cambridge).

Harman, P. M. (1982b). *Metaphysics and Natural Philosophy. The Problem of Substance in Classical Physics* (Barnes and Noble, Totowa, New Jersey).

Harrington, B. J., Park, S. Y. and Yildiz, A. (1975). 'Spectrum of heavy mesons in e^+e^- annihilation', *PRL*, **34**:168—171.

Hartle, J. B. and Hawking, S. W. (1983). 'Wave function of the universe', *PR*, **D28**:2960—2975.

Hasert, F. J. *et al.* (1973a). 'Search for elastic muon-neutrino electron scattering', *PL*, **46B**:121—124.

Hasert, F. J. *et al.* (1973b). 'Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment', *PL*, **46B**:138—140.

Hawking, S. and Ellis, G. (1968). 'The cosmic black body radiation and the existence of singularities in our universe', *Astrophys. J.*, **152**:25—36.

Hawking, S. and Penrose, R. (1970). 'The singularities of gravitational collapse and cosmology', *PRS*, **A314**:529—548.

Hawking, S. (1975). 'Particle creation by black holes', *CMP*, **43**: 199—220.

Hawking, S. (1979). 'The path-integral approach to quantum gravity', chap. 15 of *General Relativity: An Einstein Centenary Survey*, eds. Hawking, S. and Israel, W. (Cambridge University Press, Cambridge).

Hawking, S. (1980). *Is the End in Sight for Theoretical Physics?* (Cambridge University Press, Cambridge).

Hawking, S. (1987). 'Quantum cosmology', in *Three Hundred Years of Gravitation*, eds. Hawking, S. W. and Israel, W. (Cambridge, University Press, Cambridge), 631.

Hayakawa, S. (1983). 'The development of meson physics in Japan', in *The Birth of Particle Physics*, eds. Brown, L. M. and Hoddeson, L.

(Cambridge University Press, Cambridge), 82—107.
Hayashi, K. and Bregman, A. (1973). 'Poincaré gauge invariance and the dynamical role of spin in gravitational theory', *AP*, **75**:562—600.
Hehl, F. W. *et al*. (1976). 'General relativity with spin and torsion: foundations and prospects', *RMP*, **48**:393—416.
Heilbron, J. L. (1981). 'The electric field before Faraday', in *Conceptions of Ether*, eds. Cantor, G. N. and Hodge, M. J. S. (Cambridge University Press, Cambridge), 187—213.
Heisenberg, W. (1925). 'Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehung', *ZP*, **33**:879—883.
Heisenberg, W. (1926a). 'Über die spektra von Atomsystemen mit zwei Elektronen', *ZP*, **39**:499—518.
Heisenberg, W. (1926b). 'Schwankungserscheinungen und Quantenmechanik', *ZP*, **40**:501.
Heisenberg, W. (1926c). 'Quantenmechanik', *Naturwissenschaften* **14**: 989—994.
Heisenberg, W. (1926d). Letters to Pauli, W.: 28 Oct. 1926, 4 Nov. 1926, 23 Nov. 1926, in Pauli, W., *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg, u. A., Band I: 1919—1929*, eds. Hermann, A., Meyenn, K. V. and Weisskopf, V. E. (Springer-Verlag, Berlin, 1979).
Heisenberg, W. (1927). 'Über den anschaulichen Inhalt der quantentheoretischen kinematik u. mechanik', *ZP*, **43**:172—198.
Heisenberg, W. (1928). 'Zur Theorie des Ferromagnetismus', *ZP*, **49**: 619—636.
Heisenberg, W. and Pauli, W. (1929). 'Zur Quantenelektrodynamik der Wellenfelder. I', *ZP*, **56**:1—61.
Heisenberg, W. and Pauli, W. (1930). 'Zur Quantenelektrodynamik der Wellenfelder. II', *ZP*, **59**:168—190.
Heisenberg, W. (1932a). 'Über den Bau der Atomkerne', *ZP*, **77**:1—11.
Heisenberg, W. (1932b). 'Über den Bau der Atomkerne', *ZP*, **78**: 156—164.
Heisenberg, W. (1933). 'Über den Bau der Atomkerne', *ZP*, **80**: 587—596.
Heisenberg, W. (1934). 'Remerkung zur Diracschen Theorie des Positrons', *ZP*, **90**:209—231.
Heisenberg, W. (1943a). 'Die "beobachtbaren Grössen" in der Theorie der Elementarteilchen', *ZP*, **120**:513—538.
Heisenberg, W. (1943b). 'Die "beobachtbaren Grössen" in der Theorie der Elementarteilchen. II', *ZP*, **120**:673—702.
Heisenberg, W. (1944). 'Die "beobachtbaren Grössen" in der Theorie der Elementarteilchen, III', *ZP*, **123**:93—112.
Heisenberg, W. (1955). 'The development of the interpretation of the quantum theory', in *Niels Bohr and the Development of Physics*, ed. Pauli, W. (McGraw-Hill, New York), 12—29.
Heisenberg, W. (1957). 'Quantum theory of fields and elementary

particles', *RMP*, **29**:269—278.
Heisenberg, W. (1958). 'Research on the non-linear spinor theory with indefinite metric in Hilbert space', in *1958 Annual International Conference on High Energy Physics at CERN* (CERN, Geneva), 119—126.
Heisenberg, W. (1960). 'Recent research on the nonlinear spinor theory of elementary particles', in *Proceedings of the 1960 Annual International Conference on High Energy Physics at Rochester* (Interscience, New York), 851—857.
Heisenberg, W. (1961). 'Planck's discovery and the philosophical problems of atomic physics', in *On Modern Physics* (by Heisenberg and others), trans. Goodman, M. and Binns, J. W. (C. N. Potter, New York).
Heisenberg, W. (1971). *Physics and Beyond: Encounters and Conversations* (Harper and Row, New York).
Heitler, W. and Herzberg, G. (1929). *Naturwissenschaften*, **17**:673.
Heitler, W. (1936). *The Quantum Theory of Radiation* (Clarendon, Oxford).
Heitler, W. (1961). 'Physical aspects in quantum-field theory', in *The Quantum Theory of Fields*, ed. Stoops, R. (Interscience, New York), 37—60.
Hendry, J. (1984). *The Creation of Quantum Mechanics and the Bohr—Pauli Dialogue* (Reidel, Dordrecht/Boston/Lancaster).
Hertz, H. (1894). *Die Prinzipien der Mechanik in neuem Zusammenhang Dargestellt* (Barth, Leipzig).
Hesse, M. B. (1961). *Forces and Fields* (Nelson, London).
Hesse, M. B. (1974). *The Structure of Scientific Inference*, (Macmillan, London).
Hesse, M. B. (1980). *Revolutions and Reconstructions in the Philosophy of Science* (Harvester, Brighton, Sussex).
Hesse, M. B. (1981). 'The hunt for scientific reason', in *PSA 1980* (the Philosophy of Science Association, 1981) **2**:3—22.
Hesse, M. B. (1985). 'Science beyond realism and relativism', (unpublished manuscript).
Hessenberg, G. (1917). 'Vektorielle Begründung der Differentialgeometrie', *Math. Ann.*, **78**:187.
Higgs, P. W. (1964a). 'Broken symmetries, massless particles and gauge fields', *PL*, **12**:132—133.
Higgs, P. W. (1964b). 'Broken symmetries and the masses of gauge bosons', *PRL*, **13**:508—509.
Hilbert, D. (1915). 'Die Grundlagen der Physik', *Nachr. Ges. Wiss. Göttingen Math.-phys. Kl.*, 395—407.
Hilbert, D. (1917). 'Die Grundlagen der Physik: zweite Mitteilung', *Nachr. Ges. Wiss. Göttingen Math.-phys. Kl.*, 55—76, 201, 477—480.
Holton, G. (1973). *Thematic Origins of Scientific Thought: Kepler to Einstein* (Harvard University Press, Cambridge, MA).
Hosotani, Y. (1983). 'Dynamical gauge symmetry breaking as the Casimir effect', *PL*, **129B**:193—197.

Houard, J. C. and Jouvet, B. (1960). 'Etude d'un modèle de champ à constante de renormalisation nulle', *NC*, **18**:466—481.

Houtappel, R. M. F., Van Dam, H. and Wigner, E. P. (1965). 'The conceptual basis and use of the geometric invariance principles', *RMP*, **37**:595—632.

Hubble, E. P. (1929). 'A relation between distance and radial velocity among extragalactic nebulae', *Proc. Nat. Acad. of Sci. US*, **15**: 169—173.

Hurst, C. A. (1952). 'The enumeration of graphs in the Feynman—Dyson technique', *PRS*, **A214**:44—61.

Iagolnitzer, D. and Stapp, H. P. (1969). 'Macroscopic causality and physical region analyticity in the S-matrix theory', *CMP*, **14**:15—55.

Isham, C. (1990). 'An introduction to general topology and quantum topology', in *Physics, Geometry, and Topology*, ed. Lee, H. C. (Plenum, New York and London), 129—189.

Ito, D., Koba, Z. and Tomonaga, S. (1948). 'Corrections due to the reaction of "cohesive force field" to the elastic scattering of an electron. I', *PTP*, **3**:276—289.

Iwanenko, D. (1932a). 'The neutron hypothesis', *Nature*, **129**:798.

Iwanenko, D. (1932b). 'Sur la constitution des noyaux atomiques', *Compt. Rend. Acad. Sci. Paris*, **195**:439—441.

Iwanenko, D. (1934). 'Interaction of neutrons and protons', *Nature*, **133**: 981—982.

Jackiw, R. and Johnson, K. (1969). 'Anomalies of the axial-vector current', *PR*, **182**:1459—1469.

Jackiw, R. and Preparata, G. (1969). 'Probe for the constituents of the electromagnetic current and anomalous commutators', *PRL*, **22**: 975—977.

Jackiw, R., Van Royen, R., and West, G. (1970). 'Measuring light-cone singularities', *PR*, **D2**:2473—2485.

Jackiw, R. (1972). 'Field investigations in current algebra', in *Lectures on Current Algebra and Its Applications*, by Treiman, S. B., Jackiw, R. and Gross, D. J. (Princeton University Press, Princeton), 97—254.

Jackiw, R. and Rebbi, C. (1976). 'Vacuum periodicity in a Yang-Mills quantum theory', *PRL*, **37**:172—175.

Jackiw, R. (1985). 'Topological investigations of quantified gauge theories', in *Current Algebra and Anomalies*, by Treiman, S. B., Jackiw, R., Zumino, B. and Witten, E. (Princeton University Press, Princeton), 211—359.

Jackiw, R. (1991). 'Breaking of classical symmetries by quantum effects', MIT preprint CTP #1971(May 1991).

Jacob, M. (1974). *Dual Theory*(Amsterdam, North-Holland).

Jaffe, A. (1965). 'Divergence of perturbation theory for bosons', *CMP*, **1**:127—149.

Jaffe, A. (1995). Conversation with Schweber, S. S. and Cao, T. Y. on 9 February, 1995.

Jammer, M. (1974). *The Philosophy of Quantum Mechanics* (McGraw-Hill, New York).

Johnson, K. (1961). 'Solution of the equations for the Green's functions of a two dimensional relativistic field theory', *NC*, **20**:773—790.

Johnson, K. (1963). 'γ_5 Invariance', *PL*, **5**:253—254.

Johnson, K. and Low, F. (1966). 'Current algebra in a simple model', *PTP*, **37—38**:74—93.

Jordan, P. (1925). 'Über das thermische Gleichgewicht zwischen quantenatomen und Hohlraumstrahlung', *ZP*, **33**:649—655.

Jordan, P. (1927a). 'Zur quantenmechanik der gasentartung', *ZP*, **44**: 473—480.

Jordan, P. (1927b). 'Über Wellen und Korpuskeln in der quantenmechanik', *ZP*, **45**:766—775.

Jordan, P. and Klein, O. (1927). 'Zum Mehrkörperproblem der Quantentheorie', *ZP*, **45**:751—765.

Jordan, P. (1928). 'Der charakter der quantenphysik', *Naturwissenschaften*, **41**: 765—772. (English translation taken from J. Bromberg, 1976).

Jordan, P. and Wigner, E. (1928). 'Über das Paulische Äquivalenzverbot', *ZP*, **47**:631—651.

Jordan, P. (1973). 'Early years of quantum mechanics: Some reminiscences', in *The Physicist's Conception of Nature*, ed. Mehra, (Reidel, Dordrecht), 294—300.

Jost, R. (1947). 'Über die falschen Nullstellen der Eigenwerte der S-matrix', *Helv. Phys. Acta.*, **20**:256—266.

Jost, R. (1965). *The General Theory of Quantum Fields* (American Mathematical Society, Providence).

Jost, R. (1972). 'Foundation of quantum field theory', in *Aspects of Quantum Theory*, eds. Salam, A. and Wigner, E. P. (Cambridge University Press, Cambridge), 61—77.

Kadanoff, L. P. (1966). 'Scaling laws for Ising models near T_c', *Physics*, **2**:263—272.

Kadanoff, L. P., Götze, W., Hamblen, D., Hecht, R., Lewis, E., Palciauskas, V., Rayl, M., Swift, J., Aspnes, D. and Kane, J. (1967). 'Static phenomena near critical points: theory and experiment', *RMP*, **39**:395—431.

Källen, G. (1953). 'On the magnitude of the renormalization constants in quantum electrodynamics', *Dan. Mat.-Fys. Medd.*, **27**:1—18.

Källen, G. (1966). 'Review of consistency problems in quantum electrodynamics', *Acta Phys. Austr. Suppl.* **2**:133—161.

Kaluza, T. (1921). 'Zum Unitätsproblem der Physik', *Preussische Akad. Wiss. Sitzungsber.*, **54**:966—972.

Kamefuchi, S. (1951). 'Note on the direct interaction between spinor fields', *PTP*, **6**:175—181.

Kamefuchi, S. (1960). 'On Salam's equivalence theorem in vector meson theory', *NC*, **18**:691—696.

Kant, I. (1783). *Prolegomena zu einer jeden künftigen Metaphysik die als Wissenschaft wird auftreten können* (J. F. Hartknoch, Riga), English transl. revised by Ellington, J. W. (Hackett, Indianapolis, 1977).
Kastrup, H. A. (1966). 'Conformal group in space-time', *PR*, **142**: 1060—1071.
Kemmer, N. (1938). 'The charge-dependence of nuclear forces', *Proc. Cambridge Philos. Soc.*, **34**: 354—364.
Kerr, R. P. (1963). 'Gravitational field of a spinning mass as an example of algebraically special metrics', *PRL*, **11**: 237—238.
Kerszberg, P. (1989). 'The Einstein—de Sitter controversy of 1916—1917 and the rise of relativistic cosmology', in *Einstein and the History of General Relativity*, eds. Howard, D. and Stachel, J. (Birkhäuser, Boston/Basel/Berlin), 325—366.
Kibble, T. W. B. (1961). 'Lorentz invariance and the gravitational field', *JMP*, **2**: 212.
Kibble, T. W. B. (1967). 'Symmetry breaking in non-Abelian gauge theories', *PR*, **155**: 1554—1561.
Kinoshita, T. (1950). 'A note on the C meson hypothesis', *PTP*, **5**: 535—536.
Klein, F. (1872). 'A comparative review of recent researches in geometry', English trans. in *N. Y. Math. Soc. Bull.* **2**(1893): 215—249.
Klein, O. (1926). 'Quantentheorie und fündimentionale Relativitätstheorie', *ZP*, **37**: 895—906.
Klein, O. (1927). 'Zur fundimentionalen Darstellung der Relativitästheorie', *ZP*, **46**: 188.
Klein, O. (1938). *Entretiens sur les idées fondamentalles de la physique moderne* (Hermann, Paris).
Kline, M. (1972). *Mathematical Thought from Ancient to Modern Times* (Oxford University Press, Oxford and New York).
Koba, Z. and Tomonaga, S. (1947). 'Application of the "self-consistent" subtraction method to the elastic scattering of an electron', *PTP*, **2**: 218.
Koester, D., Sullivan, D. and White, D. H. (1982). 'Theory selection in particle physics: a quantitative case study of the evolution of weak-electromagnetic unification theory', *Social Studies of Science*, **12**: 73—100.
Kogut, J. and Susskind, L. (1975). 'Hamiltonian Formulation of Wilson's lattice gauge theories', *PR*, **D11**: 395—408.
Komar, A. and Salam, A. (1960). 'Renormalization problem for vector meson theories', *NP*, **21**: 624—630.
Komar, A. and Salam, A. (1962). 'Renormalization of gauge theories', *PR*, **127**: 331—334.
Koyré, A. (1965). *Newtonian Studies* (Harvard University Press, Cambridge, MA).
Krajewski, W. (1977). *Correspondence Principle and Growth of Science* (Reidel, Dordrecht).
Kramers, H. (1924). 'The law of dispersion and Bohr's theory of spectra',

Nature, **113**:673—674.
Kramers, H. and Heisenberg, W. (1925). 'Über die streuung von Strahlen durch Atome', *ZP*, **31**:681—708.
Kramers, H. (1938a). *Quantentheorie des Elektrons und der Strahlung*, part 2 of *Hand-und Jahrbuch der Chemische Physik*. *I* (Akad. Verlag, Leipzig).
Kramers, H. (1938b). 'Die Wechselwirkung zwischen geladenen Teilchen und Strahlungsfeld', *NC*, **15**:108—114.
Kramers, H. (1944). 'Fundamental difficulties of a theory of particles', *Ned. Tijdschr. Natuurk.*, **11**:134—147.
Kramers, H. (1947). A review talk at the Shelter Island conference, June 1947(unpublished). For its content and significance in the development of renormalization theory, see Schweber, S. S.: 'A short history of Shelter Island I', in *Shelter Island II*, eds. Jackiw, R., Khuri, N. N., Weinberg, S. and Witten, E. (MIT Press, Cambridge, MA, 1985).
Kretschmann, E. (1917). 'Über den Physikalischen Sinn der Relativitats-postulaten', *Ann. Phys.*, **53**:575—614.
Kripke, (1972). 'Naming and necessity', in *Semantics of Natural Language*, eds. Davidson, D. and Harman, G. (Reidel, Dordrecht).
Kronig, R. (1946). 'A supplementary condition in Heisenberg's theory of elementary particles', *Physica*, **12**:543—544.
Kruskal, M. D. (1960). 'Minimal extension of the Schwarzschild metric', *PR*, **119**: 1743—1745.
Kuhn, T. S. (1962). *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago).
Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*, second enlarged edition(University of Chicago Press, Chicago).
Kuhn, T. S. (1976). 'Theory-change as structure-change: comments on the Sneed formalism', *Erkenntnis*, **10**:179—199.
Kuhn, T. S. (1990). 'The road since structure', in *PSA 1990* (Philosophy of Science Association, East Lansing), vol. 2, 3—13.
Kuhn, T. S. (1991). 'The trouble with the historical philosophy of science', Robert and Maurine Rothschild Distinguished Lecture delivered at Harvard University on 19 November 1991.
Kuhn, T. S. (1993). Afterwords to *World Changes—Thomas Kuhn and the Nature of Science*, ed. Horwich, P. (MIT Press, Cambridge, MA).
Lamb, W. E. and Retherford, R. C. (1947). 'Fine structure of the hydrogen atom by a microwave method', *PR*, **72**:241—243.
Lagrange, J. L. (1788). *Mécanique analytique* (République, Paris).
Lagrange, J. L. (1797). *Théorie des fonctions analytiques* (République, Paris).
Landau, L. D. (1937). 'On the theory of phase transitions', *Phys. Z. Sowjetunion*, **11**:26—47, 545—555; also in *Collected Papers* (of Landau, L. D.), ed. ter Haar, D. (New York, 1965), 193—216.
Landau, L. D., Abrikosov, A. A. and Khalatnikov, I. M. (1954a). 'The removal of infinities in quantum electrodynamics', *DAN*, **95**:497—499.

Landau, L. D., Abrikosov, A. A. and Khalatnikov, I. M. (1954b). 'An asymptotic expression for the electro Green function in quantum electrodynamics', *DAN*, **95**:773—776.
Landau, L. D., Abrikosov, A. A. and Khalatnikov, I. M. (1954c). 'An asymptotic expression for the photon Green function in quantum electrodynamics', *DAN*, **95**:1117—1120.
Landau, L. D., Abrikosov, A. A. and Khalatnikov, I. M. (1954d). 'The electron mass in quantum electrodynamics', *DAN*, **96**:261—263.
Landau, L. D. (1955). 'On the quantum theory of fields', in *Niels Bohr and the Development of Physics*, ed. Pauli, W. (Pergamon, London), 52—69.
Landau, L. D. and Pomeranchuck, I. (1955). 'On point interactions in quantum electrodynamics', *DAN*, **102**:489—491.
Landau, L. D., Abrikosov, A. A. and Khalatnikov, I. M. (1956). 'On the quantum theory of fields', *NC*(Suppl.), **3**:80—104.
Landau, L. D. (1958). Letter to Heisenberg, Feb. 1958, quoted by Brown, L. M. and Rechenberg, H. (1988): 'Landau's work on quantum field theory and high energy physics (1930—1961)', Max Planck Institute Preprint MPI-PAE/Pth 42/88(July 1988), 30.
Landau, L. D. (1959). 'On analytic properties of vertex parts in quantum field theory', *NP*, **13**:181—192.
Landau, L. D. (1960a). 'On analytic properties of vertex parts in quantum field theory', in *Proceedings of the Ninth International Annual Conference on High Energy Physics(Moscow)*, II, 95—101.
Landau, L. D. (1960b). 'Fundamental problems', in *Theoretical Physics in the Twentieth Century*, eds. Fierz, M. and Weisskopf, V. F. (Interscience, New York).
Landau, L. D. (1965). *Collected Papers*(Pergamon, Oxford).
Larmor, J. (1894). 'A dynamical theory of the electric and luminiferous medium', *Philos. Trans. Roy. Soc.*, **185**:719—822.
Laudan, (1981). 'A confutation of convergent realism', *Philos. Sci.*, **48**: 19—49.
Lee, B. W. (1969). 'Renormalization of the σ-model', *NP*, **B9**: 649—672.
Lee, B. W. (1972). In *Proceedings of High Energy Physics Conference(1972)*, ed. Jacob, M. (National Accelerator Laboratory, Batavia, IL), vol. 4, 251.
Lee, T. D. (1954). 'Some special examples in renormalizable field theory', *PR*, **95**:1329—1334.
Lee, T. D. and Wick, G. C. (1974). 'Vacuum stability and vacuum excitation in a spin-0 field theory', *PR*, **D9**:2291—2316.
Lee, T. D., and Yang, C. N. (1962). 'Theory of charge vector mesons interacting with the electromagnetic field', *PR*, **128**:885—898.
Lee, Y. K., Mo, L. W. and Wu, C. S. (1963). 'Experimental test of the conserved vector current theory on the beta spectra of B^{12} and N^{12}', *PRL*, **10**:253—258.
Lehmann, H., Symanzik, K. and Zimmerman, W. (1955). 'Zur

Formalisierung quantisierter Feldtheorie', *NC*, (10)**1**:205—225.
Lehmann, H., Symanzik, K. and Zimmerman, W. (1957). 'On the formulation of quantized field theories', *NC*, (10)**6**:319—333.
Lemaître, G. (1925). 'Note on de Sitter's universe', *J. Math. Phys.* (Cambridge, Mass), **4**:189—192.
Lemaître, G. (1927). 'Un univers homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nebuleues extragalactiques', *Ann. Soc. Sci. Bruxelles*, **A47**: 49—59. [English translation: *Monthly Notices of the Royal Astronomical Society*, **91**(1931): 483—490.]
Lemaître, G. (1932). 'La expansion de l'espace', *Rev. Quest. Sci.* **20**: 391—410.
Lemaître, G. (1934). 'Evolution of the expanding universe', *Proc. Nat. Acad. Sci.*, **20**:12—17.
Lense, J. and Thirring, H. (1918). 'Ueber den Einfluss der Eigenrotation der Zentralkörper auf die Bewegung der Planeten und Monde nach der Einsteinschen Gravitationstheorie', *Phys. Z.*, **19**:156—163.
Lepage, G. P. (1989). 'What is renormalization?', Cornell University preprint CLNS 89/970, also in *From Action to Answers*, eds. DeGrand, T. and Toussaint, T. (World Scientific, Singapore, 1990), 446—457.
Leutwyler, H. and Stern, J. (1970). 'Singularities of current commutators on the light cone', *NP*, **B20**:77—101.
Levi-Civita, T. (1917). 'Nozione di parallelismo in una varietà qualunque', *Rend. Circ. Mat. Palermo*, **42**:173—205.
Lewis, G.N. (1926). 'The conservation of photons', *Nature*, **118**:874—875.
Lewis, H. W. (1948). 'On the reactive terms in quantum electrodynamics', *PR*, **73**:173—176.
Lie, S. and Engel, F. (1893). *Theorie der Transformationsgruppen*, (Teubner, Leipzig).
Lodge, O. (1883). 'The ether and its functions', *Nature*, **27**:304—306, 328—330.
London, F. (1927). 'Quantenmechanische Deutung der Theorie von Weyl', *ZP*, **42**:375—389.
Lorentz, H.A. (1892). 'La théorie électromagnetique de Maxwell et son application aux corps mouvants', *Arch. Néerlandaises*, **25**:363—552.
Lorentz, H.A. (1895). *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpen* (Brill, Leiden).
Lorentz, H.A. (1899). 'Théorie simplifiée des phénomènes électriques et optiques dans les corps en mouvement', *Versl. K. Akad. Wet. Amst.*, **7**:507—522; reprinted in *Collected Papers* (Nijhoff, The Hague, 1934—39), **5**:139—155.
Lorentz, H.A. (1904). 'Electromagnetic phenomena in a system moving with any velocity less than that of light', *Proc. K. Akad. Amsterdam*, **6**:809—830.
Lorentz, H.A. (1915). *The Theory of Electrons and its Applications to the*

Phenomena of Light and Radiant Heat, second edition (Dover, New York).
Lorentz, H. A. (1916) *Les Théories Statistiques et Thermodynamique* (Teubner, Leipzig and Berlin).
Low, F.E. (1954). 'Scattering of light of very low frequency by system of spin 1/2', *PR*, **96**:1428—1432.
Low, F. E. (1962). 'Bound states and elementary particles', *NC*, **25**: 678—684.
Low, F.E. (1967). 'Consistency of current algebra', in *Proceedings of Fifth Annual Eastern Theoretical Physics Conference*, ed. Feldman, D. (Benjamin, New York), 75.
Low, F.E. (1988). An interview at MIT, 26 July 1988.
Ludwig, G. (1968). *Wave Mechanics* (Pergamon, Oxford).
Mach, E. (1872). *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit* (Calve, Prague).
Mach, E. (1883). *Die Mechanik in ihrer Entwicklung. Historisch-Kritisch dargestellt* (Brockhaus, Leipzig).
Mack, G. (1968). 'Partially conserved dilatation current', *NP*, **B5**:499—507.
Majorana, E. (1933). 'Über die Kerntheorie', *ZP*, **82**:137—145.
Mandelstam, S. (1958). 'Determination of the pion-nucleon scattering amplitude from dispersion relations and unitarity. General theory', *PR*, **112**:1344—1360.
Mandelstam, S. (1968a). 'Feynman rules for electromagnetic and Yang-Mills fields from the gauge-independent field-theoretic formalism', *PR*, **175**:1580—1603.
Mandelstam, S. (1968b). 'Feynman rules for the gravitational field from the coordinate-independent field-theoretic formalism', *PR*, **175**:1604—1623.
Marciano, W. and Pagels, H. (1978). 'Quantum chromodynamics', *Phys. Rep.*, **36**:137—276.
Marshak, R.E. and Bethe, H.A. (1947). 'On the two-meson hypothesis', *PR*, **72**:506—509.
Matthews, P.T. (1949). 'The S-matrix for meson-nucleon interactions', *PR*, **76**:1254—1255.
Maxwell, G. (1971). 'Structural realism and the meaning of theoretical terms', in *Minnesota Studies in the Philosophy of Science*, vol. 4 (University of Minnesota Press, Minneapolis).
Maxwell, J.C. (1861/62). 'On physical lines of force', *Philos. Mag.* (4) **21**:162—175, 281—291, 338—348; **23**:12—24, 85—95.
Maxwell, J.C. (1864). 'A dynamical theory of the electromagnetic field', in *Scientific Papers*, vol. I:526—597.
Maxwell, J.C. (1873). *A Treatise on Electricity and Magnetism* (Clarendon, Oxford).
McGlinn, W.D. (1964). 'Problem of combining interaction symmetries and relativistic invariance', *PRL*, **12**:467—469.

McGuire, J. E. (1974). 'Forces, powers, aethers and fields', in *Methodological and Historical Essays in the Natural and Social Sciences*, eds. Cohen, R. S. and Wartofsky, M. W. (Reidel, Dordrecht), 119—159.

McMullin, E. (1982). 'The motive for metaphor', *Proceedings of American Catholic Philosophical Association*, **55**:27.

McMullin, E. (1984). 'A case for scientific realism', in *Scientific Realism*, ed. Leplin, J. (University of California Press, Berkeley/Los Angeles/London).

Merton, R. (1938). 'Science, technology and society in seventeenth century England', *Osiris*, **4**:360—632.

Merz, J. T. (1904). *A History of European Thought in the Nineteenth Century* (W. Blackwood, Edinburgh, 1904—1912).

Meyerson, É. (1908). *Identité et Reálité* (Labrairie Félix Alcan, Paris).

Michell, J. (1784). 'On the means of discovering the distance, magnitude, etc., of the fixed stars', *Philos. Trans. R. Soc. London*, **74**:35—57.

Mie, G. (1912a, b; 1913). 'Grundlagen einer Theorie der Materie', *Ann. Phys.*, (4)**37**:511—534; **39**:1—40; **40**:1—66.

Miller, A. (1975). 'Albert Einstein and Max Wertheimer: a gestalt psychologist's view of the genesis of special relativity theory', *History of Science*, **13**:75—103.

Mills, R. L. and Yang, C. N. (1966). 'Treatment of overlapping divergences in the photon self-energy function', *PTP* (Suppl.), **37/38**: 507—511.

Minkowski, H. (1907). 'Das Relativitätsprinzip', *Ann. Phys.*, **47**:927.

Minkowski, H. (1908). 'Die grundgleichungen für die elektromagnetischen vorguge in bewegten körper', *Gött. Nachr.*, 53—111.

Minkowski, H. (1909). 'Raum und Zeit', *Phys. Z.* **10**:104—111.

Misner, C. W., Thorne, K. S. and Wheeler, J. A. (1973). *Gravitation* (Freeman, San Francisco).

Moulines, C. U. (1984). 'Ontological reduction in the natural sciences', in *Reduction in Science*, eds. Balzer *et al.* (Reidel, Dordrecht), 51—70.

Moyer, D. F. (1978). 'Continuum mechanics and field theory: Thomson and Maxwell', *Stud. Hist. Philos. Sci.*, **9**:35—50.

Nagel, E. (1961). *The Structure of Science* (Harcourt, New York).

Nambu, Y. (1955). 'Structure of Green's functions in quantum field theory', *PR*, **100**:394—411.

Nambu, Y. (1956). 'Structure of Green's functions in quantum field theory. II', *PR*, **101**:459—467.

Nambu, Y. (1957). 'Parametric representations of general Green's functions. II', *NC*, **6**:1064—1083.

Nambu, Y. (1957a). 'Possible existence of a heavy neutral meson', *PR*, **106**:1366—1367.

Nambu, Y. (1959). Discussion remarks, *Proceedings of the International Conference on High Energy Physics*, *IX* (1959) (Academy of Science, Moscow, 1960), **2**:121—122.

Nambu, Y. (1960a). 'Dynamical theory of elementary particles suggested by superconductivity', in *Proceedings of the 1960 Annual International Conference on High Energy Physics at Rochester* (Interscience, New York), 858—866.

Nambu, Y. (1960b). 'Quasi-particles and gauge invariance in the theory of superconductivity', *PR*, **117**: 648—663.

Nambu, Y. (1960c). 'A "superconductor" model of elementary particles and its consequences', in *Proceedings of the Midwest Conference on Theoretical Physics* (Purdue University, West Lafayette).

Nambu, Y. (1960d). 'Axial vector current conservation in weak interactions', *PRL*, **4**: 380—382.

Nambu, Y. and Jona-Lasinio, G. (1961 a). 'A dynamical model of elementary particles based on an analogy with superconductivity. I', *PR*, **122**: 345—358.

Nambu, Y. and Jona-Lasinio, G. (1961b). 'A dynamical model of elementary particles based on an analogy with superconductivity. II', *PR*, **124**: 246—254.

Nambu, Y. and Lurie, D. (1962). 'Chirality conservation and soft pion production', *PR*, **125**: 1429—1436.

Nambu, Y. (1965). 'Dynamical symmetries and fundamental fields', in *Symmetry Principles at High Energies*, eds. Kursunoglu, B., Perlmutter, A. and Sakmar, A. (Freeman, San Francisco and London), 274—285.

Nambu, Y. (1989). 'Gauge invariance, vector-meson dominance, and spontaneous symmetry breaking', in Brown, Dresden, and Hoddeson (1989), 639—642.

Navier, C. L. (1821). 'Sur les lois de l'equilibre et du mouvement des corps solides elastiques', *Mém. Acad. (Paris)*, **7**: 375.

Nersessian, N. J. (1984). *Faraday to Einstein: Constructing Meaning in Scientific Theories* (Martinus Nijhoff, Dordrecht).

Neumann, C. (1868). 'Resultate einer Untersuchung über die Prinzipien der elektrodynamik', *Nachr. Ges. Wiss. Göttingen. Math.-phys. Kl.*, **20**: 223—235.

Neveu, A., and Scherk, J. (1972). 'Connection between Yang-Mills fields and dual models', *NP*, **B36**: 155—161.

Neveu, A. and Schwarz, J. H. (1971a). 'Factorizable dual model of pions', *NP*, **B31**: 86—112.

Neveu, A. and Schwarz, J. H. (1971b). 'Quark model of dual pions', *PR*, **D4**: 1109—1111.

Newman, J. R. (ed., 1946). *The Common Sense of the Exact Sciences (by W. K. Clifford)* (Simon and Schuster).

Newton, I. (1934). *Sir Isaac Newton's Mathematical Principle of Natural Philosophy and His System of the World*, ed. Cajori, F. (University of California Press, Berkeley).

Newton, I. (1978). *Unpublished Scientific Papers of Isaac Newton*, eds. Hall, A. R. and Hall, M. B. (Cambridge, University Press, Cambridge).

Nishijima, K. (1957). 'On the asymptotic conditions in quantum field

theory', *PTP*, **17**:765—802.
Nishijima, K. (1958). 'Formulation of field theories of composite particles', *PR*, **111**:995—1011.
Nissani, N. (1984). 'SL(2, C) gauge theory of gravitation: Conservation laws', *Phys. Rep.*, **109**:95—130.
Norton, J. (1984). 'How Einstein found his field equations, 1912—1915', *Hist. Stud. Phys. Sci.*, **14**:253—316.
Norton, J. (1985). 'What was Einstein's principle of equivalence?', *Stud. Hist. Philos. Sci.*, **16**:203—246.
Okubo, S. (1966). 'Impossibility of having the exact U_6 group based upon algebra of currents', *NC*, **42A**:1029—1034.
Olive, D. I. (1964). 'Exploration of S-matrix theory', *PR*, **135B**:745—760.
Oppenheimer, J.R. (1930a). 'Note on the theory of the interaction of field and matter', *PR*, **35**:461—477.
Oppenheimer, J.R. (1930b). 'On the theory of electrons and protons', *PR*, **35**:562—563.
Oppenheimer, J. R. and Snyder, H. (1939). 'On continued gravitational contraction', *PR*, **56**:455.
Oppenheimer, J.R. and Volkoff, G. (1939). 'On massive neutron cores', *PR*, **54**:540.
O'Raifeartaigh, L. (1965). 'Mass difference and Lie algebras of finite order', *PRL*, **14**:575—577.
Pais, A. (1945). 'On the theory of the electron and of the nucleon', *PR*, **68**:227—228.
Pais, A. (1966). 'Dynamical symmetry in particle physics', *RMP*, **368**:215—255.
Pais, A. and Treiman, S. (1975). 'How many charm quantum numbers are there?', *PRL*, **35**:1556—1559.
Pais, A. (1986). *Inward Bound* (Oxford University Press, Oxford).
Papapetrou, A. (1949). 'Non-symmetric stress-energy-momentum tensor and spindensity', *Philos. Mag.*, **40**:937.
Parisi, G. (1973). 'Deep inelastic scattering in a field theory with computable largemomenta behavior', *NC*, **7**:84—87.
Pasternack, S. (1938). 'Note on the fine structure of H_α and D_α', *PR*, **54**:1113—1115.
Pati, J.C. and Salam, A. (1973a). 'Unified lepton-hadron symmetry and a gauge theory of the basic interactions', *PR*, **D8**:1240—1251.
Pati, J.C. and Salam, A. (1973b). 'Is baryon number conserved?', *PRL*, **31**:661—664.
Pauli, W. (1926). Letter to Heisenberg, 19 Oct. 1926, in *Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg, U. A., Band I: 1919—1929*, eds. Hermann, A., Meyenn, K. V. and Weisskopf, V. E. (Springer-Verlag, Berlin, 1979).
Pauli, W. (1930). A letter of 4 December 1930, in *Collected Scientific Papers*, eds. Kronig, R. and Weisskopf, V.F. (Interscience, New York,

1964), Ⅱ:1313.
Pauli, W. (1933). Pauli's remarks at the seventh Solvay conference, Oct. 1933, in *Rapports du Septième Conseil de Physique Solvay*, 1933(Gauthier-Villarg, Paris, 1934).
Pauli, W. and Weisskopf, V. (1934). 'Über die quantisierung der skalaren relativistischen wellengleichung', *Helv. Phys. Acta* **7**:709—731.
Pauli, W. and Fierz, M. (1938). 'Zur Theorie der Emission langwelliger Lichtquanten', *NC*, **15**:167—188.
Pauli, W. (1941). 'Relativistic field theories of elementary particles', *RMP*, **13**:203—232.
Pauli, W. and Villars, F. (1949). 'On the invariant regularization in relativistic quantum theory', *RMP*, **21**:434—444.
Pavelle, R. (1975): 'Unphysical solutions of Yang's gravitational-field equations', *PRL*, **34**, 1114.
Peierls, R. (1934). 'The vacuum in Dirac's theory of the positive electron', *PRS*, **A146**:420—441.
Penrose, R. (1967a). 'Twistor algebra', *JMP*, **8**:345.
Penrose, R. (1967b). 'Twistor quantization and curved space-time', *Int. J. Theor. Phys.*, **1**:61—99.
Penrose, R. (1968). 'Structure of space-time', in *Lectures in Mathematics and Physics*(Battele Rencontres), eds. DeWitt, C. M. and Wheeler, J. A. (Benjamin, New York), 121—235.
Penrose, R. (1969). 'Gravitational collapse: the role of general relativity', *NC*, **1**:252—276.
Penrose, R. (1972). 'Black holes and gravitational theory', *Nature*, **236**: 377—380.
Penrose, R. (1975). 'Twistor theory', in *Quantum Gravity*, eds. Isham, C.J., Penrose, R. and Sciama, J.W. (Claredon, Oxford).
Penrose, R. (1987). In *Three Hundred Years of Gravitation*, eds. Hawking, S.W. and Israel, W. (Cambridge, University Press, Cambridge).
Perring, F. (1933). 'Neutral particles of intrinsic mass 0', *Compt. Rend. Acad. Sci. Paris*, **197**:1625—1627.
Peruzzi, I. *et al*. (1976). 'Observation of a narrow charged state at 1875 Mev/c^2 decaying to an exotic combination of K π π', *PRL*, **37**:569—571.
Peterman, A. and Stueckelberg, E. C. G. (1951). 'Restriction of possible interactions in quantum electrodynamics', *PR*, **82**:548—549.
Peterman, A. (1953a). 'Divergence. of perturbation expression', *PR*, **89**:1160—1161.
Peterman, A. (1953b). 'Renormalisation dans les séries divergentes', *Helv. Phys. Acta*, **26**:291—299.
Pickering, A. (1984). *Constructing Quarks: A Sociological History of Particle Physics*(Edinburgh University Press, Edinburgh).
Pines, D. and Schrieffer, L.R. (1958). 'Gauge invariance in the theory of superconductivity', *NC*, **10**:407—408.
Planck, M. (1900). 'Zur Theorie des Gesetzes der Energieverteilung im

Normalspectrum', *Verh. Deutsch. Phys. Ges.*, **2**:237—245.
Poincaré, H. (1890). *Electricité et optique: les théories de Maxwell et la théorie électromagnétique de la lumière*, ed. Blondin (G. Carré, Paris).
Poincaré, H. (1895). 'A propos de la theorie de M. Larmor', in *Oeuvres* (Gauthier-Villars, Paris), **9**:369—426.
Poincaré, H. (1897). 'Les idées de Hertz sur la méchanique', in *Oeuvres* (Gauthier-Villars, Paris), **9**:231—250.
Poincaré, H. (1898). 'De la mesure du temps', *Rev. Métaphys. Morale*, **6**:1—13.
Poincaré, H. (1899). 'Des fondements de la géométrie, a propos d'un livre de M. Russell', *Rev. Métaphys. Morale*, **7**:251—279.
Poincaré, H. (1900). 'La théorie de Lorentz et le principe de la réaction', *Arch. Néerlandaises*, **5**:252—278.
Poincaré, H. (1902). *La science et l'hypothèse* (Flammarion, Paris).
Poincaré, H. (1904). 'The principles of mathematical physics', in *Philosophy and Mathematics*, volume I of *Congress of Arts and Sciences: Universal Exposition, St. Louis, 1904*, ed. Rogers, H. (H. Mifflin, Boston, 1905), 604—622.
Poincaré, H. (1905). 'Sur la dynamique de i'électron', *Comptes Rendus Acad. Sci.* **140**:1504—1508.
Poincaré, H. (1906). 'Sur la dynamique de l'électron', *Rendiconti Circ. Mat. Palermo*, **21**:129—175.
Polchinski, J. (1984). 'Renormalization and effective Lagrangians', *NP*, **B231**:269—295.
Politzer, H. D. (1973). 'Reliable perturbative results for strong interactions?', *PRL*, **30**:1346—1349.
Polkinghorne, J. C. (1958). 'Renormalization of axial vector coupling', *NC*, **8**:179—180, 781.
Polkinghorne, J. C. (1967). 'Schwinger terms and the Johnson—Low model', *NC*, **52A**:351—358.
Polkinghorne, J. C. (1989). Private correspondence, 24 Oct. 1989.
Polyakov, A. M. (1974). 'Particle spectrum in quantum field theory', *JETP(Lett.)*, **20**:194—195.
Popper, K. (1970). 'A realist view of logic, physics, and history' in *Physics, Logic, and History*, eds. Yourgrau, W. and Breck, A. D. (Plenum, New York), 1—39.
Popper, K. (1974). 'Scientific reduction and the essential incompleteness of all science', in *Studies in the Philosophy of Biology, Reduction and Related Problems*, eds. Ayala, F. J. and Dobzhansky, T. (Macmillan, London), 259—284.
Post, H. R. (1971). 'Correspondence, invariance and heuristics', *Stud. Hist. Philos. Sci.*, **2**:213—255.
Putnam, H. (1975). *Mind, Language and Reality* (Cambridge University Press, Cambridge).
Putnam, H. (1978). Meaning and the Moral Sciences (Routledge and Kegan Paul, London).

Putnam, H. (1981). *Reason, Truth and History* (Cambridge University Press, Cambridge).

Radicati, L.A. (1987). 'Remarks on the early development of the notion of symmetry breaking', in *Symmetries in Physics* (1600—1980), eds. Doncel, M.G., Hermann, A., Michael, L. and Pais, A. (Bellaterra, Barcelona), 197—207.

Raine, D.J. (1975). 'Mach's principle in general relativity', *Mon. Not. R. Astron. Soc.*, **171**:507—528.

Raine, D.J. (1981). 'Mach's principle and space-time structure', *Rep. Prog. Phys.*, **44**:1151—1195.

Ramond, P. (1971a). 'Dual theory for free fermions', *PR*, **D3**:2415—2418.

Ramond, P. (1971b). 'An interpretation of dual theories', *NC*, **4A**:544—548.

Rasetti, F. (1929). 'On the Raman effect in diatomic gases. II', *Proc. Nat. Acad. Sci.*, **15**:515—519.

Rayski, J. (1948). 'On simultaneous interaction of several fields and the self-energy problem', *Acta Phys. Polonica*, **9**:129—140.

Redhead, M.L.G. (1983). 'Quantum field theory for philosophers', in *PSA 1982* (Philosophy of Science Association, East Lansing, MI, 1983), 57—99.

Regge, T. (1958a). 'Analytic properties of the scattering matrix', *NC*, **8**:671—679.

Regge, T. (1958b). 'On the analytic behavior of the eigenvalue of the S-matrix in the complex plane of the energy', *NC*, **9**:295—302.

Regge, T. (1959). 'Introduction to complex orbital momenta', *NC*, **14**:951—976.

Regge, T. (1960). 'Bound states, shadow states and Mandelstam representation', *NC*, **18**:947—956.

Reiff, J. and Veltman, M. (1969). 'Massive Yang-Mills fields', *NP*, **B13**:545—564.

Reinhardt, M. (1973). 'Mach's principle—A critical review', *Z. Naturf.* **28A**:529—537.

Resnik, M. (1981). 'Mathematics as a science of patterns: ontology and reference', *Nous*, **15**: 529—550.

Resnik, M. (1982). 'Mathematics as a science of patterns: epistemology', *Nous*, **16**:95—105.

Riazuddin, A. and Sarker, A.Q. (1968). 'Some radiative meson decay processes in current algebra', *PRL*, **20**:1455—1458.

Ricci, G. and Levi-Civita, T. (1901). 'Méthodes de calcul differentiel absolu et leurs applications', *Math. Ann.*, **54**:125—201.

Rickaysen, J.G. (1958). 'Meissner effect and gauge invariance', *PR*, **111**:817—821.

Riemann, B. (1853, 1858, 1867). In his *Gesammelte Mathematische Werke und Wissenschaftlicher Nachlass.*, ed. Weber, H. (Teubner, Leipzig, 1892).

Riemann, B. (1854). 'Über die Hypothesen, welche der Geometrie zu Grunde liegen', *Ges. Wiss. Göttingen. Abhandl.*, **13** (1867): 133—152.

Riemann, B. (1861a). In *Schwere, Elektricität, und Magnetismus: nach den Vorlesungen von Bernhard Riemann*, ed. K. Hattendoff (Hannover, 1876).

Riemann, B. (1861b). Quoted by Kline(1972), 894—896.

Rindler, W. (1956). 'Visual horizons in world models', *Mon. Not. R. Astron. Soc.*, **166**:662—677.

Rindler, W. (1977). *Essential Relativity*(Springer-Verlag, New York).

Rivier, D. and Stueckelberg, E. C. G. (1948). 'A convergent expression for the magnetic moment of the neutron', *PR*, **74**:218.

Robertson, H. P. (1939). Lecture notes, printed posthumously in *Relativity and Cosmology*, by Robertson, H. P. and Noonan, T. W. (Saunders, Philadelphia, 1968).

Rodichev, V. I. (1961). 'Twisted space and non-linear field equations', *Zh. Eksper. Ther. Fiz.*, **40**:1469.

Rohrlich, F. (1973). 'The electron: development of the first elementary particle theory', in *The Physicist's Conception of Nature*, ed. J. Mehra (Reidel, Dordrecht), 331—369.

Rohrlich, F. and Hardin, L. (1983). 'Established theories', *Philos. Sci.*, **50**:603—617.

Rosenberg, L. (1963). 'Electromagnetic interactions of neutrinos', *PR*, **129**:2786—2788.

Rosenfeld, L. (1963). Interview with Rosenfeld on 1 July 1963; Archive for the History of Quantum Physics.

Rosenfeld, L. (1968). 'The structure of quantum theory', in *Selected Papers*, eds. Cohen, R. S. and Stachel, J. J. (Reidel, Dordrecht, 1979).

Rosenfeld, L. (1973). 'The wave-particle dilemma', in *The Physicist's Conception of Nature*, ed. Mehra, J. (Reidel, Dordrecht), 251—263.

Russell, B. (1927). *The Analysis of Matter* (Allen & Unwin, London).

Sakata, S. and Tanikawa, Y. (1940). 'The spontaneous disintegration of the neutral mesotron(neutretto)', *PR*, **57**:548.

Sakata, S. and Inoue, T. (1943). First published in Japanese in *Report of Symposium on Meson Theory*, then in English, 'On the correlations between mesons and Yukawa particles', *PTP*, **1**:143—150.

Sakata, S. and Hara, O. (1947). 'The self-energy of the electron and the mass difference of nucleons', *PTP*, **2**:30—31.

Sakata, S. (1947). 'The theory of the interaction of elementary particles', *PTP*, **2**:145—147.

Sakata, S. (1950). 'On the direction of the theory of elementary particles', *Iwanami*, **II**: 100—103 (English trans. *PTP (Suppl.)*, **50** (1971):155—158.

Sakata, S., Umezawa, H. and Kamefuchi, S. (1952). 'On the structure of the interaction of the elementary particles', *PTP*, **7**:377—390.

Sakata, S. (1956). 'On the composite model for the new particles', *PTP*,

16:686—688.
Sakita, B. (1964). 'Supermultiplets of elementay particles', *PR*, **B136**: 1756—1760.
Sakurai, J.J. (1960). 'Theory of strong interactions', *AP*, **11**:1—48.
Salam, A. (1951a). 'Overlapping divergences and the S-matrix', *PR*, **82**: 217—227.
Salam, A. (1951b). 'Divergent integrals in renormalizable field theories', *PR*, **84**:426—431.
Salam, A. and Ward, J. C. (1959). 'Weak and electromagnetic interactions', *NC*, **11**:568—577.
Salam, A. (1960). 'An equivalence theorem for partially gauge-invariant vector meson interactions', *NP*, **18**:681—690.
Salam, A. and Ward, J. C. (1961). 'On a gauge theory of elementary interactions', *NC*, **19**:165—170.
Salam, A. (1962a). 'Lagrangian theory of composite particles', *NC*, **25**: 224—227.
Salam, A. (1962b). 'Renormalizability of gauge theories', *PR*, **127**:331—334.
Salam, A. and Ward, J. C. (1964). 'Electromagnetic and weak interactions', *PL*, **13**:168—171.
Salam, A. (1968). 'Weak and electromagnetic interactions', in *Elementary Particle Theory: Relativistic Group and Analyticity. Proceedings of Nobel Conference VIII*, ed. Svartholm, N. (Almqvist and Wiksell, Stockholm). 367—377.
Salam, A. and Wigner, E. P. (eds.) (1972). *Aspects of Quantum Theory* (Cambridge University Press, Cambridge).
Scadron, M. and Weinberg, S. (1964b). 'Potential theory calculations by the quasiparticle method', *PR*, **B133**:1589—1596.
Schafroth, M. R. (1951). 'Bemerkungen zur Frohlichsen theorie der supraleitung', *Helv. Phys. Acta*, **24**:645—662.
Schafroth, M.R. (1958). 'Remark on the Meissner effect', *PR*, **111**:72—74.
Scherk, J. (1975). 'An introduction to the theory of dual models and strings', *RMP*, **47**:123—164.
Schlick, M. (1918). *General Theory of Knowledge*, trans. Blumberg, A. E. and Feigl, H. (Springer-Verlag, New York).
Schmidt, B. G. (1971). 'A new definition of singular points in general relativity', *General Relativity and Gravitation*, **1**:269—280.
Schrödinger, E. (1922). 'Über eine bemerkenswerte Eigenschaft der Quantenbahnen eine einzelnen Elektrons', *ZP*, **12**:13—23.
Schrödinger, E. (1926a). 'Zur Einsteinschen Gastheorie', *Phys. Z.*, **27**: 95—101.
Schrödinger, E. (1926b). 'Quantisierung als Eigenwertproblem, Erste Mitteilung', *Ann. Phys.* **79**:361—376.
Schrödinger, E. (1926c). 'Quantisierung, Zweite Mitteilung', *Ann. Phys.* **79**:489—527.

Schrödinger, E. (1926d). 'Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen', *Ann. Phys.* **79**:734—756.

Schrödinger, E. (1926e). 'Quantisierung, Dritte Mitteilung', *Ann. Phys.* **80**:437—490.

Schrödinger, E. (1926f). 'Quantisierung, Vierte Mitteilung', *Ann. Phys.* **81**:109—139.

Schrödinger, E. (1926g). Letter to Max Planck, 31 May 1926, in *Letters On Wave Mechanics*, ed. Przibram, K., trans. Klein, M. (Philosophical Library, New York, 1967), 8—11.

Schrödinger, E. (1950). *Spacetime Structure* (Cambridge University Press, Cambridge).

Schrödinger, E. (1961). 'Wave field and particle: their theoretical relationship', 'Quantum steps and identity of particles', and 'Wave identity', in *On Modern Physics* (by Heisenberg and others), trans. Goodman, M. and Binns, J.W. (C.N. Potter, New York), 48—54.

Schwarzschild, K. (1916a). 'Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie', *Preuss. Akad. Wiss. Sitzungsber.*, 189—196.

Schwarzschild, K. (1916b). 'Über das Gravitationsfeld eines Kugel aus inkompressibler Flüssigkeit nach der Einsteinschen Theorie', *Preuss. Akad. Wiss. Sitzungsber.*, 424—434.

Schweber, S. S., Bethe, H. A. and Hoffmann, F. de. (1955). *Mesons andFields*, vol. I (Row, Peterson and Co., New York).

Schweber, S.S. (1985). 'A short history of Shelter Island I', in *Shelter Island I*, eds. Jackiw, R.; Khuri, N, N.; Weinberg, S. and Witten, E. (MIT Press, Cambridge, MA), 301—343.

Schweber, S. S. (1986). 'Shelter Island, Pocono, and Oldstone: the emergence of American quantum electrodynamics after World War II', *Osiris*, second series **2**:65—302.

Schwinger, J. (1948a). 'On quantum electrodynamics and the magnetic moment of the electron', *PR*, **73**:416—417.

Schwinger, J. (1948b). 'Quantum electrodynamics. I. A covariant formulation', *PR*, **74**:1439—1461.

Schwinger, J. (1949a). 'Quantum electrodynamics. II. Vacuum polarization and self energy', *PR*, **75**:651—679.

Schwinger, J. (1949b). 'Quantum electrodynamics. III. The electromagnetic properties of the electro-radiative corrections to scattering' *PR*, **76**:790—817.

Schwinger, J. (1951). 'On gauge invariance and vacuum polarization', *PR*, **82**:664—679.

Schwinger, J. (1957). 'A theory of the fundamental interactions', *AP*, **2**: 407—434.

Schwinger, J. (1959). 'Field theory commutators', *PRL*, **3**:296—297.

Schwinger, J. (1962a). 'Gauge invariance and mass', *PR*, **125**:397—398.

Schwinger, J. (1962b). 'Gauge invariance and mass. II', *PR*, **128**:2425—

2429.
Schwinger, J. (1966). 'Relativistic quantum field theory. Nobel lecture', *Science*, **153**:949—953.
Schwinger, J. (1967). 'Chiral dynamics', *PL*, **24B**:473—476.
Schwinger, J. (1970). *Particles, Sources and Fields* (Addison-Wesley, Reading, MA).
Schwinger, J. (1973a). *Particles. Sources and Fields*, vol. 2 (Addison-Wesley, Reading, MA).
Schwinger, J. (1973b). 'A report on quantum electrodynamics', in *The Physicist's Conception of Nature*, ed. Mehra, J. (Reidel, Dordrecht), 413—429.
Schwinger, J. (1983). 'Renormalization theory of quantum electrodynamics: an individual view', in *The Birth of Particle Physics* (eds. Brown, L. M. and Hoddeson, L. (Cambridge University Press, Cambridge), 329—375.
Sciama, D. W. (1958). 'On a non-symmetric theory of the pure gravitational field', *Proc. Cam. Philos. Soc.* **54**:72.
Seelig, C. (1954). *Albert Einstein: Eine dokumentarische Biographie* (Europa-Verlag, Zürich).
Sellars, W. (1961). 'The language of theories', in *Current Issues in the Philosophy of Science*, eds. Feigl, H. and Maxwell, G. (Holt, Rinehart, and Winston, New York), 57.
Sellars, W. (1963). 'Theoretical explanation', *Philosophy of Science*, **2**: 61.
Sellars, W. (1965). 'Scientific realism or irenic instrumentalism', in *Boston Studies in the Philosophy of Science*, eds. Cohen, R. S. and Wortofsky, M. W. (Humanities Press, New York), vol. 2:171.
Shimony, A. (1978). 'Metaphysical problems in the foundations of quantum mechanics', *International Philosophical Quarterly*, **18**:3—17.
Slater, J. C. (1924). 'Radiation and Atoms', *Nature*, 113:307—308.
Slavnov, A. A. (1972). 'Ward Identities in gauge theories', *Theor. Math. Phys.*, **19**:99—104.
Sneed, J. D. (1971). *The Logical Structure of Mathematical Physics* (Reidel, Dordrecht).
Sneed, J. D. (1979). 'Theoritization and invariance principles', in *The Logic and Epistemology of Scientific Change* (North-Holland, Amsterdam), 130—178.
Sommerfeld, A. (1911a). Letter to Planck, M. 24 April 1911.
Sommerfeld, A. (1911b). In *Die Theorie der Strahlung und der Quanten*, ed. Eucken, A. (Halle, 1913), 303.
Spector, M. (1978). *Concept of Reduction in Physical Science* (Temple University Press, Philadelphia).
Speziali, P. (ed.) (1972). *Albert Einstein-Michele Besso. Correspondance 1903—1955* (Hermann, Paris).
Stachel, J. (1980a). 'Einstein and the rigidly rotating disk', in *General Relativity and Gravitation One Hundred Years after the Birth of Albert*

Einstein, ed. Held, A. (Plenum, New York), vol. 1: 1—15.
Stachel, J. (1980b). 'Einstein's search for general covariance, 1912—1915', paper presented to the 9th International Conference on General Relativity and Gravitation, Jena, 1980; later printed in *Einstein and the History of General Relativity*, eds. Howard, D. and Stachel, J. (Birkhäuser, Boston), 63—100.

Stachel, J. (1993). 'The meaning of general covanance: the hole story', in *Philosophical Problems of the Internal and External Worlds: Essays Concerning the Philosophy of Adolf Grübaum*, eds. Janis, A. I., Rescher, N. and Massey, G.J. (University of Pittsburgh Press, Pittsburgh), 129—160.

Stachel, J. (1994). 'Changes in the concepts of space and time brought about by relativity', in *Artifacts, Representations, and Social Practice*, eds. Gould, C. C. and Cohen, R. S. (Kluwer, Dordrecht/Boston/London), 141—162.

Stapp, H.P. (1962a). 'Derivation of the CPT theorem and the connection between spin and statistics from postulated of the S-matrix theory', *PR*, **125**: 2139—2162.

Stapp, H.P. (1962b). 'Axiomatic S-matrix theory', *RMP*, **34**: 390—394.

Stapp, H. P. (1965). 'Space and time in S-matrix theory', *PR*, **B139**: 257—270.

Stapp, H.P. (1968). 'Crossing, hermitian analyticity, and the connection between spin and statistics', *JMP*, **9**: 1548—1592.

Stark, J. (1909). 'Zur experimentellen Entscheidung zwischen Ätherwellen- und Lichtquantenhypothese. I. Röntgenstrahlen', *Phys. Z.*, **10**: 902—913.

Stegmüller, W. (1976). *The Structure and Dynamics of Theories* (Springer-Verlag, New York).

Stegmüller, W. (1979). *The Structuralist View of Theories* (Springer-Verlag, New York).

Stein, H. (1981). ' "Subtle forms of matter" in the period following Maxwell', in *Conceptions of Ether*, eds. Cantor, G.N. and Hodge, M. J.S. (Cambridge University Press, Cambridge), 309—340.

Steinberger, J. (1949). 'On the use of subtraction fields and the lifetimes of some types of meson decay', *PR*, **70**: 1180—1186.

Strawson, P.T. (1950). *Mind*, **54**: 320.

Streater, R.F and Wightman, A.S. (1964). *PTC, Spin and Statistics, and All That* (Benjamin, New York).

Stueckelberg, E. C. G. (1938). 'Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kern kräfte', *Helv. Phys. Acta*, **11**: 225—244 and 299—329.

Stueckelberg, E. C. G. and Peterman, A. (1953). 'La normalisation des eonstances dans la théorie des quanta', *Helv. Phys. Acta*, **26**: 499—520.

Sudarshan, E.C.G. and Marshak, R.E. (1958). 'Chirality invariance and the universal Fermi interaction', *PR*, **109**: 1860—1862.

Sullivan, D., Koester, D., White, D. H. and Kern, K. (1980). 'Understanding rapid theoretical change in particle physics: a month-by-month co-citation analysis', *Scientometrics*, **2**:309—319.

Sutherland, D. G. (1967). 'Current algebra and some non-strong mesonic decays', *NP*, **B2**:473—440.

Symanzik, K. (1969). 'Renormalization of certain models with PCAC', (*Lett.*) *NC*, **2**:10—12.

Symanzik, K. (1970). 'Small distance behavior in field theory and power counting', *CMP*, **18**:227—246.

Symanzik, K. (1971). 'Small-distance behavior analysis and Wilson expansions', *CMP*, **23**:49—86.

Symanzik, K. (1972). In *Renormalization of Yang-Mills Fields and Applications to Particle Physics* (the Proceedings of the Marseille Conference, 19—23 June 1972), ed. Korthals-Altes, C. P. (Centre de Physique Théorique, CNRS, Marseille).

Symanzik, K. (1973). 'Infrared singularities and small-distance-behavior analysis', *CMP*, **34**:7—36.

Symanzik, K. (1983). 'Continuum limit and improved action in lattice theories', *NP*, **B226**:187—227.

Takabayasi, T. (1983). 'Some characteristic aspects of early elementary particle theory in Japan', in *The Birth of Particle Physics*, eds. Brown, L. M. and Hoddeson, L. (Cambridge University Press, Cambridge), 294—303.

Takahashi, Y. (1957). 'On the generalized Ward Identity', *NC*, **6**:371—375.

Tamm, I. (1934). 'Exchange forces between neutrons and protons, and Fermi's theory', *Nature*, **133**:981.

Tanikawa, Y. (1943). First published in Japanese in *Report of the Symposium on Meson Theory*, then in English, 'On the cosmic-ray meson and the nuclear meson', *PTP*, **2**:220.

Taylor, J.C. (1958). 'Beta decay of the pion', *PR*, **110**:1216.

Taylor, J. C. (1971). 'Ward identities and the Yang-Mills field', *NP*, **B33**:436—444.

Taylor, J. C. (1976). *Gauge Theories of Weak Interactions* (Cambridge University Press, Cambridge).

Thirring, H. (1918). 'Über die Wirkung rotierender ferner Massen in der Einsteinschen Gravitationstheorie', *Phys. Z.*, **19**:33—39.

Thirring, H. (1921). 'Berichtigung zu meiner Arbeit: "Über die Wirkung rotierender ferner Massen in der Einsteinschen Gravitationstheorie"', *Phys. Z.*, **22**:29—30.

Thirring, W. (1953). 'On the divergence of perturbation theory for quantum fields', *Heiv. Phys. Acta.* **26**:33—52.

Thomson, J.J. (1881). 'On the electric and magnetic effects produced by the motion of electrified bodies', *Philos. Mag.*, **11**:227—249.

Thomson, J.J. (1904). *Electricity and Matter* (Charles Scribner's Sons, New York).

Thomson, W. (1884). *Notes of Lectures on Molecular Dynamics and the wave Theory of Light* (Johns Hopkins University, Baltimore).

't Hooft, G. (1971a). 'Renormalization of massless Yang-Mills fields', *NP*, **B33**:173—99.

't Hooft, G. (1971b). 'Renormalizable Lagrangians for massive Yang-Mills fields', *NP*, **B35**:167—188.

't Hooft, G. (1971c). 'Prediction for neutrino-electron cross-sections in Weinberg's model of weak interactions', *PL*, **37B**:195—196.

't Hooft and Veltman, M. (1972a). 'Renormalization and regularization of gauge fields', *NP*, **B44**:189—213.

't Hooft and Veltman, M. (1972b). 'Combinatorics of gauge fields', *NP*, **B50**:318—353.

't Hooft, G. (1972c). Remarks after Symansik's presentation, in *Renormalization of Yang-Mills Fields and Applications to Particle Physics* (the Proceedings of the Marseille Conference, 19—23 June 1972), ed. Korthals-Altes, C.P. (Centre de Physique Théorique, CNRS, Marseille).

't Hooft, G. (1973). 'Dimensional regularization and the renormalization group', CERN Preprint Th. 1666, May 2, 1973, and *NP*, **B61**:455, in which his unpublished remarks at the Marseille Conference (1972c) were publicized.

't Hooft, G. (1974). 'Magnetic monopoles in unified gauge theories', *NP*, **B79**:276—284.

't Hooft, G. (1976a). 'Symmetry breaking through Bell-Jackiw anomalies', *PRL*, **37**:8—11.

't Hooft, G. (1976b). 'Computation of the quantum effects due to a four-dimensional pseudoparticle', *PR*, **D14**:3432—3450.

Toll, J. (1952). 'The dispersion relation for light and its application to problems involving electron pairs', unpublished Ph. D. dissertation, Princeton University.

Toll, J. (1956). 'Causality and the dispersion relation: logical foundations', *PR*, **104**:1760—1770.

Tomonaga, S. (1943). 'On a relativistic reformulation of quantum field theory', *Bull. IPCR* (*Rikeniko*), **22**: 545—557. (In Japanese; the English translation appeared in *PTP*, **1**:1—13).

Tomonaga, S. (1946). 'On a relativistically invariant formulation of the quantum theory of wave fields', *PTP*, **1**:27—42.

Tomonaga, S. (1948). 'On infinite reactions in quantum field theory', *PR*, **74**:224—225.

Tomonaga, S. and Koba, Z. (1948). 'On radiation reactions in collision processes. I', *PTP*, **3**:290—303.

Tomonaga, S. (1966). 'Development of quantum electrodynamics. Personal recollections', *Noble Lectures* (*Physics*): 1963—1970 (Elsevier, Amsterdam/London/New York), 126—136.

Torretti, R. (1983). *Relativity and Geometry* (Pergamon, Oxford).

Toulmin, S. (1961). *Foresight and Understanding* (Hutchinson, London).

Touschek, B.F (1957). 'The mass of the neutrino and the nonconservation

of parity', *NC*, **5**:1281—1291.
Umezawa, H., Yukawa, J. and Yamada, E. (1948). 'The problem of vacuum polarization', *PTP*, **3**:317—318.
Umezawa, H. and Kawabe, R. (1949a). 'Some general formulae relating to vacuum polarization', *PTP*, **4**:423—442.
Umezawa, H. and Kawabe, R. (1949b). 'Vacuum polarization due to various charged particles', *PTP*, **4**:443—460.
Umezawa, H. and Kamefuchi, R. (1951). 'The vacuum in quantum field theory', *PTP*, **6**:543—558.
Umezawa, H. (1952). 'On the structure of the interactions of the elementary particles, II', *PTP*, **7**:551—562.
Umezawa, H. and Kamefuchi, S. (1961). 'Equivalence theorems and renormalization problem in vector field theory (the Yang-Mills field with non-vanishing masses)' *NP*, **23**:399—429.
Utiyama, R. (1956). 'Invariant theoretical interpretation of interaction', *PR*, **101**:1597—1607.
Valatin, J.G. (1954a). 'Singularities of electron kernel functions in an external electromagnetic field', *PRS*, **A222**:93—108.
Valatin, J.G. (1954b). 'On the Dirac-Heisenberg theory of vacuum polarization', *PRS*, **A222**:228—39.
Valatin, J.G. (1954c). 'On the propagation functions of quantum electrodynamics', *PRS*, **A225**, 535—548.
Valatin, J.G. (1954d). 'On the definition of finite operator quantities in quantum electrodynamics', *PRS*, **A226**:254—265.
van Dam, H. and Veltman, M. (1970). 'Massive and massless Yang-Mills and gravitational fields', *NP*, **B22**:397—411.
van der Waerden, B. L. (1967). *Sources of Quantum Mechanics* (North-Holland, Amsterdam).
van Nieuwenhuizen, P. (1981). 'Supergravity', *PL(Rep.)*, **68**:189—398.
Vaughn, M. T., Aaron, R. and Amado, R. D. (1961). 'Elementary and composite particles', *PR*, **124**:1258—1268.
Veblen, O. and Hoffmann, D. (1930). 'Projective relativity', *PR*, **36**: 810—822.
Velo, G. and Wightman, A. (eds.) (1973). *Constructive Quantum Field Theory* (Springer-Verlag, Berlin and New York).
Veltman, M. (1963a). 'Higher order corrections to the coherent production of vector bosons in the Coulomb field of a nucleus', *Physica*, **29**:161—185.
Veltman, M. (1963b). 'Unitarity and causality in a renormalizable field theory with unstable particles', *Physica*, **29**:186—207.
Veltman, M. (1966). 'Divergence conditions and sum rules', *PRL*, **17**: 553—556.
Veltman, M. (1967). 'Theoretical aspects of high energy neutrino interactions', *PRS*, **A301**:107—112.
Veltman, M. (1968a). 'Relations between the practical results of current algebra techniques and the originating quark model', Copenhagen

Lectures, July 1968, reprinted in *Gauge Theory-Past and Future*, eds. Akhoury, R., DeWitt, B., Van Nieuwenhuizen, P. and Veltman, H. (World Scientific, Singapore, 1992).

Veltman, M. (1968b). 'Perturbation theory of massive Yang-Mills fields', *NP*, **B7**:637—650.

Veltman, M. (1969). 'Massive Yang-Mills fields', in *Proceedings of Topical Conference on Weak Interactions* (CERN Yellow Report 69—7), 391—393.

Veltman, M. (1970). 'Generalized Ward identities and Yang-Mills fields', *NP*, **B21**:288—02.

Veltman, M. (1973). 'Gauge field theories (with an appendix "Historical review and bibliography"', in the *Proceedings of the 6th International Symposium on Electron and Photon Interactions at High Energies*, eds. Rollnik, H. and Pheil, W. (North-Holland, Amsterdam).

Veltman, M. (1977). 'Large Higgs mass and μ-e universality', *PL*, **70B**: 253—254.

Veltman, M. (1991). Interview with the author, 9—11 Dec. 1991.

Veltman, M. (1992). 'The path to renormalizability', a talk given at the *Third International Symposium on the History of Particle Physics*, *SLAC*, 24—27 June 1992.

Waller, I. (1930). 'Bemerküngen über die Rolle der Eigenenergie des Elektrons in der Quantentheorie der Strählung', *ZP*, **62**:673—676.

Ward, J.C. (1950). 'An identity in quantum electrodynamics', *PR*, **78**: 182.

Ward, J. C. (1951). 'On the renormalization of quantum electrodynamics', *Proc. Phys. Soc. London*. **A64**:54—6.

Ward, J. C. and Salam. A. (1959). 'Weak and electromagnetic interaction', *NC*, **11**:568—577.

Weinberg, S. (1960). 'High energy behavior in quantum field theory', *PR*, **118**:838—849.

Weinberg, S. (1964a). 'Systematic solution of multiparticle scattering problems', *PR*, **B133**:232—256.

Weinberg, S. (1964b). 'Derivation of gauge invariance and the equivalence principle from Lorentz invariance of the S-matrix', *PL*, **9**:357—359.

Weinberg, S. (1964c). 'Photons and gravitons in S-matrix theory: derivation of charge conservation and equality of gravitational and inertial mass', *PR*, **B135**:1049—1056.

Weinberg, S. (1965a). In *Lectures on Particles and Field Theories*, eds. Deser, S. and Ford, K.W. (Prentice/Englewood, NJ).

Weinberg, S. (1965b). 'Photons and gravitons in perturbation theory: derivation of Maxwell's and Einstein's equations', *PR*, **B138**:988—1002.

Weinberg, S. (1967a). 'Dynamical approach to current algebra', *PRL*, **18**:188—191.

Weinberg, S. (1967b). 'A model of leptons', *PRL*, **19**:1264—66.

Weinberg, S. (1968). 'Nonlinear realizations of chiral symmetry', *PR*, **166**:1568—1577.

Weinberg, S. (1977). 'The search for unity: notes for a history of quantum field theory', *Daedalus*, **106**:17—35.
Weinberg, S. (1978). 'Critical phenomena for field theorists', in *Understanding the Fundamental Constituents of Matter*, ed. Zichichi, A. (Plenum, New York), 1—52.
Weinberg, S. (1979). 'Phenomenological Lagrangian', *Physica*, **96A**: 327—340.
Weinberg, S. (1980a). 'Conceptual foundations of the unified theory of weak and electromagnetic interactions', *RMP*, **52**:515—523.
Weinberg, S. (1980b). 'Effective gauge theories', *PL*, **91B**:51—55.
Weinberg, S. (1983). 'Why the renormalization group is a good thing', in *Asymptotic Realms of Physics: Essays in Honor of Francis. E. Low*, eds. Guth, A. H.; Huang, K. and Jaffe, R. L. (MIT Press, Cambridge, MA), 1—19.
Weinberg, S. (1985). 'The ultimate structure of matter', in *A Passion for Physics: Essays in Honor of Jeffrey Chew*, eds. De Tar, C., Finkelstein, J. and Tan, C.I. (Taylor & Francis, Philadelphia), 114—127.
Weinberg, S. (1986a). 'Particle physics: past and future', *Int. J. Mod. Phys*. **A1/1**:135—145.
Weinberg, S. (1986b). 'Particles, fields, and now strings', in *The Lesson of Quantum Theory*, eds. De Boer, J., Dal, E. and Ulfbeck, O. (Elsevier, Amsterdam), 227—239.
Weinberg, S. (1987). 'Newtonianism, reductionism and the art of congressional testimony', in *Three Hundred Years of Gravitation*, eds. Hawking, S. W. and Israel, W. (Cambridge, University Press, Cambridge).
Weinberg, S. (1992). *Dreams of a Final Theory* (Pantheon).
Weinberg, S. (1995). *The Quantum Theory of Fields* (Cambridge University Press, Cambridge).
Weisberger, W. I. (1965) 'Renormalization of the weak axial-vector coupling constant', *PRL*, **14**:1047—1055.
Weisberger, W. I. (1966). 'Unsubstracted dispersion relations and the renormalization of the weak axial-vector coupling constants', *PR*, **143**: 1302—1309.
Weisskopf, V.F (1934). 'Über die Selbstenergie des Elektrons', *ZP*, **89**: 27—39.
Weisskopf, V. F. (1936). 'Über die Elektrodynamic des Vakuums auf Grund der Quantentheorie des Elektrons', *K. Danske Vidensk. Selsk., Mat.-Fys. Medd*. **14**:1—39.
Weisskopf, V.F. (1939). 'On the self-energy and the electromagnetic field of the electron', *PR*, **56**:72—85.
Weisskopf, V. F. (1949). 'Recent developments in the theory of the electron', *RMP*, **21**:305—328.
Weisskopf, V. F. (1972). *Physics in the Twentieth Century* (MIT Press, Cambridge, MA).
Weisskopf, V.F. (1983). 'Growing up with field theory: the development

of quantum electrodynamics', in *The Birth of Particle Physics*, eds. Brown, L. M. and Hoddeson, L. (Cambridge University Press, Cambridge), 56—81.

Welton, T. A. (1948). 'Some observable effects of the quantum-mechanical fluctuations of the electromagnetic field', *PR*, **74**:1157—1167.

Wentzel, G. (1933a, b; 1934). 'Über die Eigenkräfie der Elementarteilchen. I, II and III', *ZP*, **86**:479—494, 635—645; **87**: 726—733.

Wentzel, G. (1943). *Einführung in der Quantentheorie der Wellenfelder* (Franz Deuticke, Vienna); English translation: *Quantum Theory of Fields*, (Interscience, New York, 1949).

Wentzel, G. (1956). 'Discussion remark', in *Proceedings of 6th Rochester Conference on High Energy Nuclear Physics* (New York), VIII-15 to VIII-17.

Wentzel, G. (1958). 'Meissner effect', *PR*, **112**:1488—1492.

Wentzel, G. (1959). 'Problem of gauge invariance in the theory of the Meissner effect', *PRL*, **2**:33—34.

Wentzel, G. (1960). 'Quantum theory of fields (until 1947)', in *Theoretical Physics in the Twentieth Century: A Memorial Volume to Wolfgang Pauli* (Interscience, New York), 44—67.

Wess, J. (1960). 'The conformal invariance in quantum field theory', *NC*, **13**:1086.

Wess, J. and Zumino, B. (1971). 'Consequences of anomalous Ward identities', *PL*, **37B**:95—97.

Wess, J. and Zumino, B. (1974). 'Supergauge transformations in four dimensions', *NP*, **B70**:39—50.

Weyl, H. (1918a). *Raum-Zeit-Materie* (Springer, Berlin).

Weyl, H. (1918b). 'Gravitation und Elektrizität', *Preussische. Akad. Wiss. Sitzungsber.*, 465—478.

Weyl, H. (1921). 'Electricity and gravitation', *Nature*, **106**: 800—802.

Weyl, H. (1922). 'Gravitation and electricity', in *The Principle of Relativity* (Methuen, 1923), 201—216.

Weyl, H. (1929). 'Elektron und Gravitation. I', *ZP*, **56**:330—352.

Weyl, H. (1930). *Gruppentheorie und Quantenmechanik* (translated by Robertson as *The Theory of Groups and Quantum Mechanics* (E. P. Dutton, New York, 1931).

Wheeler, J. (1962). *Geometrodynamics* (Academic, New York).

Wheeler, J. (1964a). 'Mach's principle as a boundary condition for Einstein's equations', in *Gravitation and Relativity*, eds. Chiu, H.-Y. and Hoffman, W. F. (Benjamin, New York), 303—349.

Wheeler, J. (1964b). 'Geometrodynamics and the issue of the final state', in *Relativity, Group and Topology*, eds. DeWitt, C. and DeWitt, B. (Gordon and Breach, New York), 315.

Wheeler, J. (1973). 'From relativity to mutability', in *The Physicist's Conception of Nature*, ed. Mehra, J. (Reidel, Dordrecht), 202—247.

Whewell, W. (1847). *The Philosophy of the Inductive Sciences*, second

edition(Parker & Son, London).
Whittaker, E. (1951). *A History of the Theories of Aether and Electricity* (Nelson, London).
Widom, B. (1965a). 'Surface tension and molecular correlations near the critical point', *J. Chem. Phys.*, **43**:3892—3897.
Widom, B. (1965b). 'Equation of state in the neighborhood of the critical point', *J. Chem. Phys.*, **43**:3898—3905.
Wien, W. (1909). *Encykl. der Math. Wiss*(Leipzig), vol. 3:356.
Wightman, A.S. (1956). 'Quantum field theories in terms of expectation values', *PR*, **101**:860—866.
Wightman, A.S. (1976). 'Hilbert's sixth problem: mathematical treatment of the axioms of physics', in *Mathematical Developments Arising from Hilbert Problems*, ed. Browder, F.E. (American Mathematical Society, Providence), 147—240.
Wightman, A.S. (1978). 'Field theory, Axiomatic', in the *Encyclopedia of Physics*(McGraw-Hill, New York), 318—321.
Wightman, A.S. (1986). 'Some lessons of renormalization theory', in *The Lesson of Quantum Theory*, eds. de Boer, J., Dal, E. and Ulfbeck, D. (Elsevier, Amsterdam), 201—225.
Wightman, A.S. (1989). 'The general theory of quantized fields in the 1950s', in *Pions to Quarks*, eds. Brown, L.M., Dresden, M. and Hoddeson, L. (Cambridge University Press, Cambridge), 608—629.
Wightman, A.S. (1992). Interview.
Wigner, E.P. and Witmer, E.E. (1928). 'Über die Struktur der zweiatomigen Molekulspektren nach der Quantenmechanik', *ZP*, **51**: 859—886.
Wigner, E.P. (1931). *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*(Braunschweig).
Wigner, E. (1937). 'On the consequences of the symmetry of the nuclear Hamiltonian on the spectroscopy of nuclei', *PR*, **51**:106—119.
Wigner, E. (1949). 'Invariance in physical theory', in *Proc. Amer Philos. Soc.*, **93**:521—526.
Wigner, E. (1960). 'The unreasonable effectiveness of mathematics in the natural sciences', *Commun. Pure. Appl. Math.* **13**:1—14.
Wigner, E. (1963). Interview with Wigner on 4 Dec 1963; Archive for the History of Quantum Physics.
Wigner, E. (1964a). 'Symmetry and conservation laws', in *Proc. Nat. Acad. Sci.*, **5**:956—965.
Wigner, E. (1964b). 'The role of invariance principles in natural philosophy', in *Proceedings of the Enrico Fermi International School of Physics*(XXIX), ed. Wigner, E. (New York), ix—xvi.
Wilson, K.G. (1965). 'Model Hamiltonians for local quantum field theory', *PR*, **B140**:445—457.
Wilson, K.G. (1969). 'Non-Lagrangian models of current algebra', *PR*, **179**:1499—1512.
Wilson, K.G. (1970a). 'Model of coupling-constant renormalization',

PR, **D2**:1438—1472.
Wilson, K. G. (1970b). 'Operator-product expansions and anomalous dimensions in the Thirring model', *PR*, **D2**:1473—1477.
Wilson, K.G.(1970c). 'Anomalous dimensions and the breakdown of scale invariance in perturbation theory', *PR*, **D2**:1478—1493.
Wilson, K. G. (1971a). 'The renormalization group and strong interactions', *PR*, **D3**:1818—1846.
Wilson, K. G. (1971b). 'Renormalization group and critical phenomena. I. Renormalization group and the Kadanof scaling picture', *PR*, **B4**: 3174—3183.
Wilson, K. G. (1971c). 'Renormalization group and critical phenomena. II. Phasespace cell analysis of critical behavior', *PR*, **B4**:3184—3205.
Wilson, K. G. (1972). 'Renormalization of a scalar field in strong coupling', *PR*, **D6**:419—426.
Wilson, K. G. and Fisher, M. E. (1972). 'Critical exponents in 3.99 dimensions', *PRL*, **28**:240—243.
Wilson, K.G. (1974). 'Confinement of quarks', *PR*, **D10**:2445—2459.
Wilson, K.G. and Kogut, J. (1974). 'The renormalization group and the ε expansion', *PL(Rep.)*, **12C**:75—199.
Wilson, K.G. (1975). 'The renormalization group: critical phenomena and the Kondo problem', *RMP*, **47**:773—840.
Wilson, K.G. (1979). 'Problems in physics with many scales of length', *Scientific American*, **241**(2):140—157.
Wilson, K. G. (1983). 'The renormalization group and critical phenomena', *RMP*, **55**:583—600.
Wise, N. (1981). 'German concepts of force, energy and the electromagnetic ether: 1845—1880', in Contor and Hodge(1981), 269—308.
Wu, T.T. and Yang, C.N. (1967). 'Some solutions of the classical isotopic gauge field equations', in *Properties of Matter under Unusual Conditions*, eds. Mark, H. and Fernbach, S. (Wiley, New York), 349—354.
Wu, T. T. and Yang, C. N. (1975). 'Concept of nonintegrable phase factors and global formulation of gauge fields', *PR*, **D12**:3845—3857.
Yang, C. N. (1970). 'Charge quantization of the gauge group, and flux quantization', *PR*, **D1**:8.
Yang, C.N. (1974). 'Integral formalism for gauge fields', *PRL*, **33**:445—447.
Yang, C. N. (1977). 'Magnetic monopoles, fiber bundles, and gauge fields', *Ann. N. Y. Acad. Sci.*, **294**:86—97.
Yang, C. N. (1980). 'Einstein's impact on theoretical physics', *Physics Today*, **33**(6):42—49.
Yang, C.N. (1983). *Selected Papers(1945—1980)*(Freeman, New York).
Yang, C.N. and Mills, R. L. (1954a). 'Isotopic spin conservation and a generalized gauge invariance', *PR*, **95**:631.
Yang, C.N. and Mills, R. L. (1954b). 'Conservation of isotopic spin and isotopic gauge invariance', *PR*, **96**:191—195.

Yates, F. (1964). *Giordano Bruno and the Hermetic Tradition* (Routledge and Kegan Paul, London).
Yoshida, R. (1977). *Reduction in the Physical Sciences*. (Dalhousie University Press, Halifax).
Yoshida, R. (1981). 'Reduction as replacement', *Brit. J. Philos. Sci.*, **32**:400.
Yukawa, H. (1935). 'On the interaction of elementary particles', *Proc. Phys.-Math. Soc. Japan*, **17**:48—57.
Yukawa, H. and Sakata, S. (1935a). 'On the theory of internal pair production', *Proc. Phys.-Math. Soc. Japan*, **17**:397—407.
Yukawa, H. and Sakata, S. (1935b). 'On the theory of the β-disintegration and the allied phenomenon', *Proc. Phys.-Math. Soc. Japan*, **17**:467—479.
Zachariasen, F. and Zemach, C. (1962). 'Pion resonances', *PR*, **128**: 849—858.
Zahar, E. (1982). 'Poincaré et la découverte du Principe de Relativité', LSE-Preprint, 6 Feb. 1982.
Zilsel, E. (1942). 'The sociological roots of science', *Amer. J. Sociology*, **47**:544—562.
Zimmermann, W. (1958). 'On the bound state problem in quantum field theory', *NC*, **10**:597—614.
Zimmermann, W. (1967). 'Local field equation for A^4-coupling in renormalized perturbation theory', *CMP*, **6**:161—188.
Zumino, B. (1975). 'Supersymmetry and the vacuum', *NP*, **B89**: 535—546.
Zweig, G. (1964a). 'An SU_3 model for strong interaction symmetry and its breaking: I and II', CERN preprint TH401(17 Jan. 1964) and TH 412 (21 Feb. 1964).
Zweig, G. (1964b). 'Fractionally charged particles and SU(6)', in *Symmetries in Elementary Particle Physics*, ed. Zichichi, A. (Academic Press, New York), 192—243.
Zwicky, F. (1935). 'Stellar guests', *Scientific Monthly*, **40**:461.
Zwicky, F. (1939). 'On the theory and observation of highly collapsed stars', *PR*, **55**:726—743.

人 名 索 引

注：此处页码是原书的页码、中译本的边码。下同。——译者

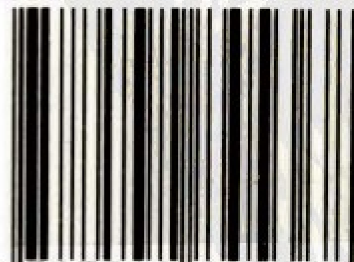


主 题 索 引



上架建议：科普读物

ISBN 978-7-5428-4536-8



定价：55.00 元
易文网：www.ewen.cc